STATISTICAL INFERENCE FOR NORMAL MEANS WITH ORDER RESTRICTIONS AND APPLICATIONS TO DOSE-RESPONSE STUDIES

KARELYN ALEXANDREA DAVIS

Statistical Inference for Normal Means with Order Restrictions and Applications to Dose-Response Studies

by

© *Karelyn Alexandrea Davis*

A thesis submitted to the

School of Graduate Studies

in partial fulfillment of the

requirements for the degree of

Master of Science

Department of Mathematics & Statistics

Memorial University of Newfoundland

May 2004

St. John's **Newfoundland** **Canada**

Abstract

Scientific experiments often compare several treatment means with a control mean. In particular, such multiple comparisons arise in biopharmaceutical studies in which it is desirable to conduct the inferences in a specified order and failure to achieve the desired inference at any step renders subsequent comparisons unnecessary. In clinical trials, an important dosing quantity is the minimum effective dose (MED), defined as the minimum dose such that the mean response is clinically significantly better than the mean response of the control by a practical significant difference. In relation to MED estimation, previous authors have either failed to account for the monotonicity of the dose-response means or considered the case of a zero clinically significant difference. In this thesis, an innovative approach using Kuhn-Tucker conditions to evaluate the optimal confidence lower bound at each step in a closed step-down testing procedure is derived and simulation results are presented.

Acknowledgments

"A proof is a proof. What kind of proof? It's a proof. And when you have good proof, it's because it's proven." - The Rt. Hon. Jean Chrétien

The author would like to gratefully acknowledge the encouragement and advice of Dr. Chu-In Charles Lee of the Department of Mathematics and Statistics, Memorial University of Newfoundland, throughout the preparation of this thesis.

The present research was supported by graduate fellowships from the Natural Sciences and Engineering Research Council of Canada and the School of Graduate Studies, Memorial University of Newfoundland.

In addition, I would like to recognize my parents and brother Mark for always extending their love and positive support in all my endeavours.

Table of Contents

List of Tables

List of Appendices

Chapter 1

Introduction

In the past half-century, statisticians have recognized the improvement in efficiency of many inference problems as a result of implementing the prior ordering of parameters or restrictions in the analysis. Problems of this type may originate from diverse areas of study: an educator may wish to determine if levels of distraction varying from none to excessive during an examination result in scores in the reverse order of magnitude; a sociologist may examine if people in low, middle and high socioeconomic groups possess low, middle and high knowledge of certain current events; and a National Hockey League (NHL) owner may be interested in determining whether selecting players with a high ranking in the Entry Draft will lead to improved team performance. (Daniel, 1990 and

Dawson and Magee, 2001) Alternative hypothesis of this nature are referred to as ordered alternatives and are studied in the general area of order restricted statistical inference.

Furthermore, the focus of many scientific experiments details the comparison of several treatment means with a control mean. When a treatment is significantly better than the control, researchers wish to evaluate the difference between the best treatment and the control. For example, suppose $\mu = (\mu_1, \ldots, \mu_k)$, a vector of mean effects of k treatments, where μ_1 is the mean of the control and $\mu_2, \ldots, \mu_k$ are the mean responses corresponding to increasing dose of a test drug. Then, in one-sided comparisons with the control where a significant difference is of interest, as in dose-response studies, the desired inferences are $\mu_i > \mu_1 + \delta$, where δ defines a practical significant difference.

One instance of such multiple comparisons occurs when it is desirable to give the inferences in a specified order and failure to achieve the desired inference at any step renders subsequent comparisons unnecessary. This situation arises in dose-response and toxicity studies, where $\mu_2, \ldots, \mu_k$ correspond to increasing dose of a substance. In dose-response studies, it is desirable that a method not declare a lower dose to be efficacious if it does not declare a higher dose to be efficacious.

The pharmaceutical industry has implemented numerous order restricted inferences throughout the development of a drug. While the development process may have multiple objectives, the establishment of a therapeutic window, or range of effective doses, is of considerable interest. In particular, a practical dosing quantity used in biopharmaceutical studies is known as the minimum effective dose (MED), which is defined as the minimum dose such that the mean response at that dose is significantly better than the mean response of the controls (Ruberg (1995a,b), Hsu and Berger (1999)). Determination of the MED may involve hypothesis testing, regression methods, or a combination of both.

Since dose-response means increase for increasing dose, we require the assumption of monotonicity of the μ_i:

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k \tag{1.1}$$

Moreover, for the MED problem, Bauer (1997) showed that only the pairwise contrasts between the ith response mean and the control mean strongly control the type I familywise error rate (FWE) in a stepwise testing procedure, regardless of whether the above assumption of monotonicity is satisfied. If condition (1.1) is not satisfied, other procedures may lead to excessive error rates. However, as pairwise contrasts do not exploit any prior knowledge of the shape of the dose

response function. they are less powerful.

Previous research has considered likelihood ratio tests (LRTs) and multiple comparison tests in a stepwise procedure. Simulation studies conducted by various authors (e.g. Ruberg (1989), Tamhane, Hochberg and Dunnett (1996), Dunnett and Tamhane (1998), Hsu and Berger (1999) and Liu (2001)) have shown that those procedures which account for the monotonicity of the response means are the most powerful. However, with the exception of Hsu and Berger (1999), such analyses have assumed the clinically significant difference (δ) to be zero, which is not true in general.

With respect to statistical inference, a confidence interval provides a visual perspective superior to a point estimate or a test statistic. The problem of obtaining confidence intervals under ordered restrictions has received mild recognition by researchers, primarily due to the general intractability of these types of problems (p. 405 of Robertson, Wright and Dykstra (1988)). In a recent paper, Hsu and Berger (1999) proposed a stepwise confidence set method, however this method did not assume monotonicity of the response means.

As stated in Dunnett and Tamhane (1998), the problem of identifying the MED is formulated as a sequence of hypothesis testing problems, beginning with a comparison of the largest dose versus the control dose and continuing

in a stepwise fashion. When the null hypothesis is rejected in favour of the alternative hypothesis at any step, there exists at least one treatment better than the control. With the monotonicity assumption (1.1), we note that $\mu_k - \mu_1$ is the largest difference between any treatment mean and the control mean, hence the confidence lower bound for this difference will be bounded below by that for any $\mu_i - \mu_1$ $(i = 2, \ldots, k)$ or any resulting non-negative linear combinations. We state that μ_k is significantly larger than μ_1 if the maximized confidence lower bound for the difference in means is larger than the clinically significant difference δ, and thus reject the null hypothesis. However, the likelihood ratio test cannot be used to provide confidence intervals. In this thesis, a detailed construction of the simultaneous confidence lower bound for $\mu_k - \mu_1$ is discussed, which is a particularly useful inference method not previously considered in relation to this problem.

The procedure outlined in this paper will be to devise a Theorem to calculate the optimal lower bound, as noted above. Marcus and Peritz (1976) found the optimal lower bound by calculating the lower bound over all partitions and hence selecting the maximum. This lengthy search is simplified by using Kuhn–Tucker conditions to retrieve the optimal coefficients directly for the given data set. For a slightly different problem, Lee, Peng and Liu (2002) derived algorithms

to compute the optimal difference between treatment and control means when treatments are at least as good as the control and when no restriction is placed on the response means.

The present thesis considers both known and innovative results to explain the nature of simultaneous confidence lower bounds and their applications in dose-response studies. Chapter 2 provides technical results useful for further work including definitions of the maximum likelihood estimate under ordered restrictions, tests of simply ordered hypothesis, Kuhn-Tucker conditions and derivations of a multiple contrast test statistic and simultaneous confidence lower bounds. Chapter 3 details the application of this problem to ordered ANOVA by Marcus and Peritz (1976), and the derivation of an optimization theorem for the aforementioned difference. The optimization theorem leads to an efficient algorithm, for which calculations are illustrated by way of a numerical example at the end of the chapter. In Chapter 4, various aspects of dose-response studies are discussed including a definition of the MED and procedures for its identification. A subsequent numerical example examines computational aspects of the various procedures, including the new procedure developed in Chapter 3. Other approaches available in recent literature for determining the MED are presented in the last section of Chapter 4. In Chapter 5, the design and results of an

in-depth simulation study which examines the power efficiency of the MED procedures are provided. Finally, a summary of results obtained and suggestions for future work are given in Chapter 6.

Chapter 2

Technical Results

The experiment to be considered in this thesis is a model for ANOVA with ordered restrictions. Denote a set of increasing dose levels by 1, 2, ..., k where 1 corresponds to the zero or control dose level. A one-way model is discussed, in which n_i experimental units are tested at the ith dose level, $i = 1, \ldots, k$. Let observations Y_{ij} be mutually independent with $Y_{ij} \sim N(\mu_i, \sigma^2)$, $j = 1, \ldots, n_i$ and $i = 1, \ldots, k$. Then $\bar{Y}_i \sim N(\mu_i, \sigma^2/n_i)$, $i = 1, \ldots, k$ are the sample means of the dose groups and let $S^2 = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2/\nu$ be an unbiased estimate of the common variance σ^2, with $\nu = \sum_{i=1}^{k} n_i - k > 0$. Then S^2 is distributed as $\sigma^2 \chi^2_\nu/\nu$, independently of $\bar{Y}_1, \ldots, \bar{Y}_k$. The parameter space for this problem is defined as $\Omega = \{\mu \in R^k : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k\}$, with σ^2 as a nuisance parameter.

2.1 Tests of Simply Ordered Hypothesis

2.1.1 Maximum Likelihood Estimate Under Order Restrictions

The restricted maximum likelihood estimator of μ subject to Ω is denoted $\mu^\star = (\mu_1^\star, \ldots, \mu_k^\star)$ and is denoted the isotonic regression of $\bar{\mathbf{Y}} = (\bar{Y}_1, \ldots, \bar{Y}_k)$ under Ω with weights $n_1, \ldots, n_k$.

As the data is assumed to be normally distributed, the maximum likelihood estimate (MLE) is the solution to the following constrained weighted least squares problem:

$$\min \sum_{i=1}^{k} n_i(\bar{Y}_i - \mu_i)^2 \qquad \text{such that} \quad \mu \in \Omega. \tag{2.1}$$

The MLE is readily calculated using the Pool-Adjacent-Violators Algorithm (PAVA) (see Robertson, Wright and Dykstra (1988)). The process is essentially a successive averaging of $\bar{Y}_i$'s until a sequence of non–decreasing values is obtained. The MLE of the μ's may then be partitioned into consecutive sequences of equal–valued $\mu^\star$'s such as

$$\mu_1^\star = \cdots = \mu_{i_1}^\star \;<\; \mu_{i_1+1}^\star = \cdots = \mu_{i_2}^\star \;<\; \cdots \;<\; \mu_{i_{l-1}+1}^\star = \cdots = \mu_k^\star, \tag{2.2}$$

with $i_0 = 0$ and $i_l = k$.

With the previous representation, the following results are valid.

Lemma 2.1.1. The vector $\mu^\star$ is the MLE of μ if and only if

$$\begin{aligned} \sum_{i=1}^{k} n_i(\bar{Y}_i - \mu_i^\star)\mu_i^\star &= 0 \qquad \text{and} \\ \sum_{i=1}^{k} n_i(\bar{Y}_i - \mu_i^\star)\nu_i &\leq 0 \qquad \text{for all } \nu \in \Omega. \end{aligned}$$

Lemma 2.1.2. With $\mu^\star$ as the MLE of μ, then for $r = 1, \ldots, l$,

$$\mu^\star_{i_{r-1}+1} = \cdots = \mu^\star_{i_r} = \frac{\sum_{i=i_{r-1}+1}^{i_r} n_i \bar{Y}_i}{\sum_{i=i_{r-1}+1}^{i_r} n_i} = A_r. \tag{2.3}$$

Lemma 2.1.3. With $\mu^\star$ as the MLE of μ, if

$$\mu^\star_{i_{r-1}+1} = \cdots = \mu^\star_{i_r} = A_r < \mu^\star_{i_r+1} \implies \frac{\sum_{i=j}^{i_r} n_i \bar{Y}_i}{\sum_{i=j}^{i_r} n_i} \leq A_r.$$

for $j = 1, \ldots, i_r$.

2.1.2 Likelihood Ratio Test

As is often the case in applications, a researcher may believe that the response means are monotone increasing, *a priori*, thus likelihood ratio tests (LRTs) for homogeneity of normal means with simple order restrictions are introduced. As

previously noted, the interest of this paper lies with a variation of the following hypothesis under the monotonicity assumption $\mu_1 \leq \cdots \leq \mu_k$. The LRT for ordered alternatives was introduced by Bartholomew (1959a,b, 1961a,b) and further discussed by Robertson, Wright and Dykstra (1988) as follows:

$$H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$$

$$H_1 : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$$

The LRT rejects H_0 in favour of H_1 for large values of the test statistic

$$S_{01} = \frac{\sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2}{\sum_{i=1}^{k} n_i(\bar{Y}_i - \hat{\mu})^2/\nu + S^2},$$

where $\hat{\mu} = \sum_{i=1}^{k} n_i\bar{Y}_i / \sum_{i=1}^{k} n_i$. When σ^2 is known, the test statistic is given by

$$\bar{\chi}^2_{01} = \frac{\sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2}{\sigma^2}.$$

As shown in Robertson, Wright and Dykstra (1988), as $\nu \to \infty$, the distribution of S_{01} approaches that of $\bar{\chi}^2_{01}$.

The null distributions of S_{01} and $\bar{\chi}^2_{01}$ under H_0 are

$$P[S_{01} \geq s] = \sum_{j=2}^{k} P_s(j, k; \mathbf{w}) P[F_{j-1,N-j} \geq \frac{s(N-j)}{\nu(j-1)}] \qquad (2.4)$$

$$P[\bar{\chi}^2_{01} \geq s] \;=\; \sum_{j=2}^{k} P_s(j,k;\mathbf{w}) P[\chi^2_{j-1} \geq s] \tag{2.5}$$

for any $s > 0$, where $N = \sum_{i=1}^{k} n_i$, $\mathbf{w} = (n_1, \ldots, n_k)$. $P_s(j,k;\mathbf{w})$ is the level probability that $\mu^\star$ takes j distinct values under H_0 and χ^2_{j-1} is a chi–squared variable with $j-1$ degrees of freedom. For the case of equal weights, the level probabilities and the critical values for S_{01} and $\bar{\chi}^2_{01}$ are tabled in Robertson, Wright and Dykstra (1988). We now discuss the calculation of level probabilities in more detail.

For the simply ordered case, i.e. with $\mu_1 \leq \cdots \leq \mu_k$, the level probabilities are denoted $P_s(l,k;\mathbf{w})$. When $k = 2$, the level probabilities are $P_s(1,2;\mathbf{w}) = P_s(2,2;\mathbf{w}) = \frac{1}{2}$. If the weights are equal, the level probabilities are more readily obtained. For this case, we omit the weights from the notation and denote the level probabilities as $P_s(l,k)$. It is demonstrated by Robertson, Wright and Dykstra (1988) that the $P_s(l,k)$ are distribution free over the collection of independent, identically distributed continuous random variables, i.e. the probability that the isotonic regression of $Y_1, Y_2, \ldots, Y_k$ with a simple order and equal weights has l level sets does not depend in the distribution of the Y_i, provided they are independent with a common continuous distribution. Furthermore, an expression for the probability generating function (PGF) of $\{P_s(l,k)\}$ is obtained and

used to derive a recurrence relationship for the equal–weights level probabilities as follows:

Theorem 2.1.4. Robertson, Wright and Dykstra (1988).

The probabilities $P_s(l,k)$ satisfy

$$P_s(1,k) = \frac{1}{k} \qquad \text{and} \qquad P_s(k,k) = \frac{1}{k!}$$

and

$$P_s(l,k) = \frac{1}{k}P_s(l-1,k-1) + \frac{k-1}{k}P_s(l,k-1)$$

for $l = 2, 3. \ldots, k-1$.

Moreover, Hogg (1965) noted the relationship between the likelihood ratio function and the class of linear functions of the sample mean $\bar{Y}_i$. Without loss of generality, assume that $\sum_{i=1}^{k} n_i c_i = 0$ and $\sum_{i=1}^{k} n_i c_i^2 = 1$ for the linear contrast $\sum_{i=1}^{k} n_i c_i \bar{Y}_i$ and the k populations have equal known variance σ^2. With the further assumption that $\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k$, the following result by Hogg (1965) is presented:

Theorem 2.1.5. Hogg (1965).

$$\sqrt{\bar{\chi}_{01}^2} \;=\; \max\{n_i c_i \bar{Y}_i / (\sigma^2 \sum_{i=1}^{k} n_i c_i^2)^{1/2}\},$$

subject to c_i satisfies the order $c_1 \leq c_2 \leq \cdots \leq c_k$. The maximum is attained at

$c_j^\star = (\mu_j^\star - \hat{\mu})/\sqrt{\sum_{i=1}^k n_i(\mu_i^\star - \hat{\mu})^2}$, $j = 1, \ldots, k$.

Since the linear contrast is normally distributed, the power function is more easily computed for the test based on linear contrasts than $\bar{\chi}_{01}^2$. Furthermore, a linear contrast may be decomposed into the sum of independent, normally distributed variables which is particularly useful if the hypothesis that $\mu_1 = \cdots = \mu_k$ is resolved into a number of nested hypothesis (Hogg, 1965).

2.2 Multiple Contrast Test Statistic

When $S_{01} > s_{k,\nu,\alpha}$, one rejects H_0 and concludes that treatment mean μ_k is significantly larger than μ_1. However, there is no corresponding simultaneous confidence lower bound for $\mu_k - \mu_1$ when $k > 2$. The following test statistic is introduced:

$$T = \max_{\mathbf{c} \in \mathbf{C}} \sum_{i=1}^k n_i c_i \bar{Y}_i / S(\sum_{i=1}^k n_i c_i^2)^{1/2}, \tag{2.6}$$

where $\mathbf{C} = \{\mathbf{c} = (c_1, c_2, \ldots, c_k) : \sum_{i=1}^k n_i c_i = 0, c_1 \leq c_2 \leq \cdots \leq c_k\}$.

Let $t_{k,\nu,\alpha}$ be the critical value of statistic T, which leads to

$$P_\mu\{\sum_{i=1}^k n_i c_i \mu_i \geq \sum_{i=1}^k n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S(\sum_{i=1}^k n_i c_i^2)^{1/2}, \text{for all } \mathbf{c} \in \mathbf{C}\} = \mathbf{1} - \alpha. \tag{2.7}$$

We may rewrite the left–hand side of the above as

$$P_\mu\{\max_{\mathbf{c}\in\mathbf{C}} \sum_{i=1}^{k} n_i c_i(\bar{Y}_i - \mu_i)/S(\sum_{i=1}^{k} n_i c_i^2)^{1/2} \leq t_{k,\nu,\alpha}, \mu \in \Omega\}$$

$$= P_0\{\max_{\mathbf{c}\in\mathbf{C}} \sum_{i=1}^{k} n_i c_i \bar{Y}_i/S(\sum_{i=1}^{k} n_i c_i^2)^{1/2} \leq t_{k,\nu,\alpha}\}$$

$$= P_0\{\sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2/S^2 \leq t_{k,\nu,\alpha}^2\}$$

where the last equation is the result of an argument similar to one in Hogg (1965). Therefore, we write

$$T^2 = \sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2/S^2 \tag{2.8}$$

- **Distribution of T^2 under $\mu_1 = \cdots = \mu_k$**

As stated by Lee, Peng and Liu (2002), the right-hand side of the previous equation is given by Wright (1988) for a different purpose. The statistic T^2 is asymptotically equivalent to S_{01}. The null distribution of T under $\mu_1 = \cdots = \mu_k$ may be written as:

$$P_0(T \geq t) = \sum_{j=1}^{k} P(j,k,\mathbf{w}) P[F_{j-1,\nu} \geq \frac{t^2}{j-1}], \tag{2.9}$$

for any $t > 0$, where $P(j,k,\mathbf{w})$ is the level probability under $\mu_1 = \cdots = \mu_k$ that $\mu^\star$ takes j distinct values and $\mathbf{w} = (n_1, n_2, \ldots, n_k)$. The critical value $t_{k,\nu,\alpha}$ is the value t when one equates (2.9) to α.

2.3 Simultaneous Confidence Bounds

We define the set $\mathbf{C}$ which places restrictions on the "scores" represented by $\mathbf{c} = (c_1, c_2, \ldots, c_k)$:

$$\mathbf{C} \;=\; \{\mathbf{c} = (c_1, c_2, \ldots, c_k) : \sum_{i=1}^{k} n_i c_i = 0, \quad c_i \leq c_{i+1}, i = 1, \ldots, k-1\}.$$

For the ordered ANOVA model, Marcus and Peritz (1976) state the following lemma as the basis for the one-sided simultaneous confidence bound:

Lemma 2.3.1 (Marcus and Peritz (1976)).

$$\mathrm{P}_{\mu}\{\sum_{i=1}^{k} n_i c_i \mu_i > \sum_{i=1}^{k} n_i c_i \bar{Y}_i - d_{\alpha}(\sum_{i=1}^{k} n_i c_i^2)^{1/2}, \quad \forall \mathbf{c} \in \mathbf{C}\} \;=\; 1 - \alpha,$$

where d_{α} is the upper 100α percentile of the null distribution of the square root of Bartholomew's likelihood ratio statistic when $\sigma^2 = 1$.

Lemma 2.3.1 defines simultaneous confidence bounds for certain classes of linear functions of expectations. These bounds also hold without the restriction of monotonicity on the parameter space. Taking these restrictions into account, the bounds may be improved in certain cases without altering the confidence level of $1 - \alpha$.

According to Lemma 2.3.1, the $1 - \alpha$ one–sided simultaneous confidence

bound for any contrast $\sum_{i=1}^{k} n_i c_i \mu_i$ with $c_i \leq c_{i+1}, i = 1, 2. \ldots, k-1$ is given by

$$l(\sum_{i=1}^{k} n_i c_i \mu_i) \;=\; \sum_{i=1}^{k} n_i c_i \bar{Y}_i - d_\alpha (\sum_{i=1}^{k} n_i c_i^2)^{1/2}. \tag{2.10}$$

For the general case, as considered by Marcus and Peritz (1976), the following set is now defined:

$$\tau_{\mathbf{c}^\star} = \{\mathbf{c} : \mathbf{c} \in \mathbf{C}, \sum_{i=1}^{k} n_i c_i \mu_i \leq \sum_{i=1}^{k} n_i c_i^\star \mu_i, \quad \forall \mu : \mu_i \leq \mu_{i+1}, i = 1, \ldots, k-1\} \tag{2.11}$$

for a given $\mathbf{c}^\star = (c_1^\star, \ldots, c_k^\star)$.

The improved confidence lower bound is denoted by

$$S(\tau_{\mathbf{c}^\star}) \;=\; \max_{\mathbf{c} \in \tau_{\mathbf{c}^\star}} l(\sum_{i=1}^{k} n_i c_i \mu_i) \;=\; \max_{\mathbf{c} \in \tau_{\mathbf{c}^\star}} \{\sum_{i=1}^{k} n_i c_i \bar{Y}_i - d_\alpha (\sum_{i=1}^{k} n_i c_i^2)^{1/2}\}. \tag{2.12}$$

One now has the following lemma:

Lemma 2.3.2 (Marcus and Peritz (1976)). With μ monotone nondecreasing,

$$\mathrm{P}_\mu \{\sum_{i=1}^{k} n_i c_i \mu_i > S(\tau_{\mathbf{c}^\star}), \quad \forall \mathbf{c} : c_i \leq c_{i+1}, \sum_{i=1}^{k} n_i c_i = 0\} \;=\; 1 - \alpha.$$

The condition $\sum_{i=1}^{k} n_i c_i \mu_i \leq \sum_{i=1}^{k} n_i c_i^\star \mu_i$ for all nondecreasing sequences of μ_i is equivalent to

$$\sum_{i=j}^{k} n_j c_j \leq \sum_{i=j}^{k} n_j c_j^\star \quad \text{for all } i = 2, \ldots, k. \tag{2.13}$$

For the step-down confidence set test procedure, often used in dose-response studies, we need only maximize the one–sided confidence lower bound for the contrast $\mu_k - \mu_1$. In a similar manner as Marcus and Peritz (1976), equation (2.7) implies that a $1-\alpha$ simultaneous confidence lower bound for any contrast $\sum_{i=1}^{k} n_i c_i \mu_i$ with $c_1 \leq c_2 \leq \cdots \leq c_k$ is given by

$$l(\sum_{i=1}^{k} n_i c_i \mu_i) \;=\; \sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i c_i^2)^{1/2}. \tag{2.14}$$

In particular, the $1-\alpha$ simultaneous confidence lower bound for the difference $\mu_i - \mu_1$ (i.e. the difference between the ith treatment mean μ_i and the control mean μ_1) is

$$l(\mu_i - \mu_1) \;=\; \bar{Y}_i - \bar{Y}_1 - t_{k,\nu,\alpha} S (n_i^{-1} + n_0^{-1})^{1/2}.$$

To determine the minimum effective dose (MED), we implement a step–down procedure which maximizes the above one–sided confidence lower bound for the contrast $\mu_k - \mu_1$. The set $\mathcal{K}$ is now defined as

$$\mathcal{K} \;=\; \{\mathbf{c} : \sum_{i=1}^{k} n_i c_i = 0,\; c_1 \leq c_2 \cdots \leq c_k,\; \sum_{i=1}^{k} n_i c_i \mu_i \leq \mu_k - \mu_1,\; \mu \in \Omega\}.$$

Thus, the improved confidence lower bound for $\mu_k - \mu_1$ is denoted by

$$L(\mu_k - \mu_1) \;=\; \max_{\mathbf{c} \in \mathcal{K}} l(\sum_{i=1}^{k} n_i c_i \mu_i) \;=\; \max_{\mathbf{c} \in \mathcal{K}} \{\sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i c_i^2)^{1/2}\}. \tag{2.15}$$

In Chapter 3, an Optimization Theorem and an Iterative Algorithm are derived to calculate the optimal lower bound, $L(\mu_k - \mu_1)$ for a given sample. Chapter 4 examines the implementation of a step-wise testing procedure for finding the minimum effective dose (MED) by calculating $L(\mu_k - \mu_1)$ for dose levels $2, \ldots, k$.

2.4 Kuhn-Tucker Conditions

The evaluation of the improved simultaneous confidence lower bounds such as in (2.15) is a maximization problem subject to equality and inequality constraints. Let $\mathbf{x}$ be an $n \times 1$ vector and $H(\mathbf{x})$ be an $m \times 1$ vector whose components $h_1(\mathbf{x}), \ldots, h_m(\mathbf{x})$ are differentiable concave functions of $\mathbf{x} \geq 0$. In addition, let $g(\mathbf{x})$ be a differentiable concave function of $\mathbf{x} \geq 0$. The Kuhn-Tucker equivalence theorem will determine an $\mathbf{x}^o$ that maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq 0$ and $\mathbf{x} \geq 0$. A vector $\mathbf{x}$ is said to be feasible if it satisfies all given constraints. The optimal value of the problem is the maximum of $g(\mathbf{x})$ over the sets of feasible points. Those feasible points which attain the optimal value are called optimal solutions. Let $\left[\frac{\partial \varphi}{\partial x_i}\right]^o$ and $\left[\frac{\partial \varphi}{\partial v_j}\right]^o$ denote the partial derivatives evaluated at a particular point $\mathbf{x}^o$ and $\mathbf{v}^o$ respectively.

Theorem 2.4.1 (Equivalence Theorem). Let $h_1(\mathbf{x}), \ldots, h_m(\mathbf{x}), g(\mathbf{x})$ be concave as well as differentiable for $\mathbf{x} \geq 0$. Let $\phi(\mathbf{x}, \mathbf{v}) = g(\mathbf{x}) + \mathbf{v}'H(\mathbf{x})$. Then $\mathbf{x}^o$ is a solution that maximizes $g(\mathbf{x})$ constrained by $H(\mathbf{x}) \geq 0$ and $\mathbf{x} \geq 0$ if and only if $\mathbf{x}^o$ and some $\mathbf{v}^o$ satisfy the following conditions:

$$(1) \qquad \left[\frac{\partial \phi}{\partial x_i}\right]^o \leq 0, \quad \left[\frac{\partial \phi}{\partial x_i}\right]^{o'} x^o = 0, \quad \mathbf{x^o} \geq 0;$$

$$(2) \qquad \left[\frac{\partial \phi}{\partial v_j}\right]^o \geq 0, \quad \left[\frac{\partial \phi}{\partial v_j}\right]^{o'} v^o = 0, \quad \mathbf{v^o} \geq 0.$$

(Theorem 3 Kuhn-Tucker 1951)

Simple modifications are made when the constraints $H(\mathbf{x}) \geq 0$, $\mathbf{x} \geq 0$ are changed to the following three cases:

Case 1. $H(\mathbf{x}) \geq 0$.

Here, using $\phi(\mathbf{x}, \mathbf{v}) = g(\mathbf{x}) + \mathbf{v}'H(\mathbf{x})$ defined for all $\mathbf{x}$ and constrained only by $\mathbf{v} \geq 0$, one must replace condition (1) by

$$(1^\star) \qquad \left[\frac{\partial \phi}{\partial x_i}\right]^o = 0.$$

Case 2. $H(\mathbf{x}) = 0$, $\mathbf{x} \geq 0$.

In this case, using $\phi(\mathbf{x}, \mathbf{v}) = g(\mathbf{x}) + \mathbf{v}'H(\mathbf{x})$ defined for all $\mathbf{v}$ and constrained only by $\mathbf{x} \geq 0$, one must replace condition (2) by

$$(2^\star) \qquad \left[\frac{\partial \phi}{\partial u_j}\right]^o = 0.$$

Case 3. $H(\mathbf{x}) = 0$.

Here, using $\phi(\mathbf{x}, \mathbf{v}) = g(\mathbf{x}) + \mathbf{v}'H(\mathbf{x})$ defined for all $\mathbf{x}$ and $\mathbf{v}$ without constraints, one must replace conditions (1) and (2) by (1*) and (2*). This corresponds to the familiar method of Lagrange multipliers.

Chapter 3

Optimization Theorem

3.1 Preliminary Results

3.1.1 Application to Ordered ANOVA by Marcus and Peritz (1976)

In their 1976 article, Marcus and Peritz noted that likelihood ratio statistics for hypothesis testing in certain restricted normal models have been available for some years, but no corresponding simultaneous confidence (SC) procedures have been presented. Simultaneous confidence lower bounds were obtained and an application to analysis of variance (ANOVA) was presented for any contrast

of the form $\sum_{i=1}^{k} n_i c_i \mu_i$. The results presented in this section are for this general case, however examples are stated which use the Marcus and Peritz method for the specific contrast $\mu_k - \mu_1$.

In summary, to find the optimal lower bound, $L(\sum n_i c_i \mu_i)$, as defined by equation (2.12) and discussed in Section 2.3, the solution obtained by Marcus and Peritz (1976) involved the following steps:

Firstly, let $l(\sum n_i c_i \mu_i) = \sum_{i=1}^{k} n_i c_i \bar{Y}_i - d_\alpha (\sum_{i=1}^{k} n_i c_i^2)^{1/2}$, hence $L(\sum n_i c_i \mu_i) = \max_{\mathbf{c} \in \tau_{\mathbf{c}^\star}} [l(\sum n_i c_i \mu_i)]$.

After amalgamating the sample means and obtaining the nondecreasing maximum likelihood estimators, the authors note that $L(\sum n_i c_i \mu_i)$ is nonnegative for all $\mathbf{c}^\star$, and thus consider only positive-valued $l(\sum n_i c_i \mu_i)$'s.

Furthermore, suppose that the maximum of $l(\sum n_i c_i \mu_i)$ in $\tau_{\mathbf{c}^\star}$ is obtained at a point $\mathbf{c_0}$ such that

$$\sum_{j=1}^{i} n_j c_{0j} = \sum_{j=1}^{i} n_j c_j^\star \quad \text{for } i = i_1, \ldots, i_m \text{ and no other values of } i.$$

Using the fact that $\sum_{i=1}^{k} n_i c_i = \sum_{i=1}^{k} n_i c_i^\star = 0$, then alternatively,

$$\sum_{i \in R_s} n_i c_{0i} = \sum_{i \in R_s} n_i c_i^\star \quad \text{for } s = 1, \ldots, m, \tag{3.1}$$

where $R_1 = \{1, \ldots, i_1\}, \ldots, R_s = \{i_{s-1} + 1, \ldots, i_s\}, \ldots, R_m = \{i_{m-1} + 1, \ldots, k\}$.

After implementing Lagrange multipliers, the authors obtain:

$$c_{0i} = \bar{c}_s^\star + (\mu_i^\star - \bar{X}_s)(\sum_s n_s \bar{c}_s^{\star 2})^{1/2} \Bigg/ \left\{ d_\alpha^2 - \sum_s \sum_{i \in R_s} n_i (\mu_i^\star - \bar{X}_s)^2 \right\}^{1/2} \tag{3.2}$$

and

$$l(\sum n_i c_{0i} \mu_i) \;=\; \sum_s n_s \bar{c}_s^\star \bar{X}_s - (\sum_s n_s \bar{c}_s^{\star 2})^{1/2} \left\{ d_\alpha^2 - \sum_s \sum_{i \in R_s} n_i (\mu_i^\star - \bar{X}_s)^2 \right\}^{1/2} \tag{3.3}$$

where $n_s = \sum_{i \in R_s} n_i$, $\bar{c}_s^\star = n_s^{-1} \sum_{i \in R_s} n_i c_i^\star$ and $\bar{X}_s = n_s^{-1} \sum_{i \in R_s} n_i \mu_i^\star$. Equation (3.3) requires that

$$d_\alpha^2 \;>\; \sum_s \sum_{i \in R_s} n_i (\mu_i^\star - \bar{X}_s)^2. \tag{3.4}$$

To find the optimal lower bound, let $P = \{R_0, \ldots, R_m\}$ denote any partition of $\{1, \ldots, k\}$ and let $\mathcal{P}$ be the set of all partitions satisfying equation (3.4) for which the right hand side of (3.2) satisfies (2.13). Then, the optimal lower bound $L(\sum n_i c_i \mu_i) = \max_{P \in \mathcal{P}} l(\sum n_i c_{0i} \mu_i)$.

Marcus and Peritz remark that some partitions may be eliminated. For instance, letting partition P_1 be a refinement of partition P_0 (i.e. all the sets in partition P_1 are subsets of the sets in P_0), then (i) if P_0 satisfies equation (3.4) then so does P_1 and (ii) $l(\sum n_i c_{0i} \mu_i)_{P_1} \;\leq\; l(\sum n_i c_{0i} \mu_i)_{P_0}$. Moreover, $\mathcal{P}$ is not empty as it always includes $P : R_i = \{i\} (i = 1, \ldots, k)$. Also, for any $P \in \mathcal{P}$, the right hand side of equation (3.2) is increasing in i (and hence the vector $\mathbf{c}$

corresponding to P is always in $\tau_{\mathbf{c}^\star}$). If $i, i+1 \in R_s$, $c_i \leq c_{i+1}$ follows directly from (3.2): if $i \in R_s$ and $i+1 \in R_{s+1}$ one has, by Lemma 2.3.2 and equation (3.1) that $c_i \leq c_i^\star$ and $c_{i+1} \geq c_{i+1}^\star$ and thus $c_i \leq c_{i+1}$.

In summary, the approach of Marcus and Peritz (1976) calculates the optimal bounds for various partitions and then finds the maximum of these bounds. For large values of k, it is apparent that the method becomes tedious. For example, Section 3.2 of Davis (2002) calculates the confidence lower bound for $\mu_4 - \mu_1$ for a given data set, i.e. when $k = 4$. The method by Marcus and Peritz necessitates several pages of computations. thus a more efficient method should be found.

3.1.2 Preparatory Lemmas for Optimization Theorem

To implement a step-wise procedure for determining the minimum effective dose, an exact method is needed to calculate the optimal lower bound $L(\mu_k - \mu_1)$. As noted in the previous section. Marcus and Peritz (1976) calculated the optimal bound for numerous partitions and selected the maximum, however the method is inefficient. This section demonstrates the steps required to derive a theorem which establishes and necessary and sufficient condition for the optimal solution.

The following lemmas rewrite the conditions outlined in the set $\mathcal{K}$. We note that the statement and proof of the following lemmas are similar to those of Peng

(2002, Chapter 7). In the latter paper, the author considered the assumption of $\mu_q \leq \mu_{q+1} \leq \cdots \leq \mu_k$, with $q \geq 2$, and alternate definitions of the sets $\mathbf{C}$ and $\mathcal{K}$. A multiple contrast test statistic and simultaneous confidence lower bounds for $\mu_k - \mu_1$ were also derived under this more general assumption. In this paper, we will furthermore consider the parameter space $\Omega = \{\mu : \mu_1 \leq \mu_2 \leq \cdots \leq \mu_k\}$.

Lemma 3.1.1. For $\mu \in \Omega$ and $\mathbf{c} \in \mathbf{C}$, then $\sum_{i=1}^{k} n_i c_i \mu_i \leq \mu_k - \mu_1$ if and only if $\sum_{i=j}^{k} n_i c_i \leq 1$, for $j = 2, \ldots, k$.

Proof of Lemma 3.1.1.

($\Longleftarrow$) Since $\sum_{i=1}^{k} n_i c_i = 0$, $c_1 \leq \cdots \leq c_k$, then $c_1 \leq 0$, $c_k \geq 0$ and there exists a j such that $j > 1$ with $c_j > 0$ and $c_{j-1} \leq 0$. (i.e. j divides c_i's into positive and negative components). Therefore,

$$\sum_{i=1}^{k} n_i c_i = \sum_{i=j}^{k} n_i c_i - \sum_{i=1}^{j-1} n_i |c_i| = 0.$$

If $\sum_{i=j}^{k} n_i c_i \leq 1$ then

$$\begin{aligned} \sum_{i=1}^{k} n_i c_i \mu_i &\leq \left(\sum_{i=j}^{k} n_i c_i \right) \mu_k - \left(\sum_{i=1}^{j-1} n_i |c_i| \right) \mu_1 \\ &= \left(\sum_{i=j}^{k} n_i c_i \right) (\mu_k - \mu_1) \qquad \text{from above} \\ &\leq \mu_k - \mu_1, \qquad \text{as required.} \end{aligned}$$

($\Longrightarrow$) Given $c_1 \leq \cdots \leq c_k$, $\sum n_i c_i = 0$, suppose that $\sum_{i=1}^{k} n_i c_i \mu_i \leq \mu_k - \mu_1$ for all $\mu \in \Omega$. Let $\mu_i = 1$ for $i = j, \ldots, k$ and $\mu_i = 0$ for $i = 1, \ldots, j-1$. Then,

$$\sum_{i=1}^{k} n_i c_i \mu_i = \sum_{i=j}^{k} n_i c_i \leq \mu_k - \mu_1 = 1 - 0 = 1,$$

which proves the result. $\square$

Lemma 3.1.2. Let $\mu^\star$ be the MLE of μ under Ω. If $\mu_i^\star = \mu_{i+1}^\star$ then $c_i^o = c_{i+1}^o$, where $\mathbf{c^o}$ is the optimal solution to (2.15).

Proof of Lemma 3.1.2.

As before, we represent the monotonicity of the MLE as follows;

$$\mu_{i_{r-1}}^\star < \mu_{i_{r-1}+1}^\star = \cdots = \mu_{i_r}^\star < \mu_{i_r+1}^\star.$$

with $i_{r-1} + 1 \leq i < i_r$. Suppose $c_i^o < c_{i+1}^o$. We shall show it leads to a contradiction. Thus, define a new value b_i as

$$\begin{aligned} b_i &= c_i^o \qquad \forall\, i \neq i_{r-1}+1, \ldots, i_r \\ b_i &= \frac{\displaystyle\sum_{j=i_{r-1}+1}^{i_r} n_j c_j^o}{\displaystyle\sum_{j=i_{r-1}+1}^{i_r} n_j} \qquad i = i_{r-1}+1, \ldots, i_r \end{aligned}$$

Note that $c_{i_{r-1}+1}^o \leq b_i \leq c_{i_r}^o$. By parts (b) and (c) of Lemma 4 of Marcus and

Peritz (1976), the following results are valid:

$$\begin{cases} \sum_{i=1}^{k} n_i c_i^o \bar{Y}_i \leq \sum_{i=1}^{k} n_i b_i \bar{Y}_i \\ \sum_{i=1}^{k} n_i c_i^{o2} \geq \sum_{i=1}^{k} n_i b_i^2 \end{cases}$$

Therefore,

$$\sum_{i=1}^{k} n_i c_i^o \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i c_i^{o2})^{1/2} \leq \sum_{i=1}^{k} n_i b_i \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i b_i^2)^{1/2},$$

which is impossible, since $\mathbf{c^o}$ is the optimal solution. Hence, $c_i^o = c_{i+1}^o$. □

When $\mu_i^\star = \mu_{i+1}^\star$ we may combine the treatments as a single treatment with a total sample size $n_i + n_{i+1}$.

Using Lemmas 3.1.1 and 3.1.2, we may rewrite the problem of equation (2.15) into the following problem:

Lemma 3.1.3. The optimal solution to (2.15) is equivalent to the solution to the following problem:

$$\max \left\{ \sum_{i=1}^{k} n_i c_i \mu_i^\star - t_{k,\nu,\alpha} S \sqrt{\sum_{i=1}^{k} n_i c_i^2} \right\} \tag{3.5}$$

subject to $\mathbf{c} \in \mathbf{C}$ and $\sum_{i=j}^{k} n_i c_i \leq 1$.

Proof of Lemma 3.1.3.

Analogous to Peng (2002), to show the equivalence of equations (2.15) and (3.5), we must prove they have the same solution and the same value. For convenience, let $f(\mathbf{c})$ represent equation (2.15) and $g(\mathbf{c})$ represent equation (3.5).

Let $\mathbf{c^o}$ be the optimal solution to (3.5). Then, $g(\mathbf{c}) \leq g(\mathbf{c^o})$ for any $\mathbf{c} \in \mathcal{K}$. From equation (2.2), Lemma 2.1.2 and Lemma 3.1.2, we have the following:

$$\sum_{i=i_{r-1}+1}^{i_r} n_i c_i^o \bar{Y}_i = c_{i_1}^o \sum_{i=i_{r-1}+1}^{i_r} n_i \bar{Y}_i = c_{i_1}^o \sum_{i=i_{r-1}+1}^{i_r} n_i A_r = \sum_{i=i_{r-1}+1}^{i_r} n_i c_i^o \mu_i^\star.$$

It follows that $f(\mathbf{c^o}) = g(\mathbf{c^o}) \geq g(\mathbf{c})$. From the second inequality of Lemma 2.1.1 and since $\mathbf{c} \in \Omega$, let $\nu = \mathbf{c}$ thus

$$\sum_{i=1}^{k} n_i c_i (\bar{Y}_i - \mu_i^\star) \leq 0 \implies \sum_{i=1}^{k} n_i c_i \bar{Y}_i \leq \sum_{i=1}^{k} n_i c_i \mu_i^\star$$

as required. It follows that $g(\mathbf{c}) \geq f(\mathbf{c})$. $\square$

The following lemma relates the optimal solution to the MLE.

Lemma 3.1.4. If $\mathbf{c^o}$ is the optimal solution to (3.5) subject to $\sum_{i=1}^{k} n_i c_i = 0$ and $\sum_{i=1}^{k} n_i c_i \mu_i \leq \mu_k - \mu_1$, then

$$c_1^o \leq c_2^o \leq \cdots \leq c_k^o.$$

Proof of Lemma 3.1.4.

We wish to show if $\mu_i^\star < \mu_{i+1}^\star \Longrightarrow c_i^o \leq c_{i+1}^o$, $i \geq l$. Suppose $\mu_i^\star < \mu_{i+1}^\star$ and $c_i^o > c_{i+1}^o$. This implies that $(c_i^o - c_{i+1}^o)(\mu_{i+1}^\star - \mu_i^\star) > 0$. Define a new constant γ such that

$$\begin{aligned} \gamma_j &= c_j^o \qquad j \neq i, i+1 \\ \gamma &= \frac{n_i c_i^o + n_{i+1} c_{i+1}^o}{n_i + n_{i+1}} \qquad j = i, i+1 \end{aligned}$$

Therefore, if $\mathbf{c^o}$ is an optimal solution to (3.5) then we have

$$A + n_i c_i^o \mu_i^\star + n_{i+1} c_{i+1}^o \mu_{i+1}^\star - t_{k,\nu,\alpha} S \sqrt{B + n_i c_i^{o2} + n_{i+1} c_{i+1}^{o2}} \qquad \text{(a)}$$

$$A + n_i \gamma \mu_i^\star + n_{i+1} \gamma \mu_{i+1}^\star - t_{k,\nu,\alpha} S \sqrt{B + n_i \gamma^2 + n_{i+1} \gamma^2} \qquad \text{(b)}$$

where

$$A = \sum_{j \neq i,\ i+1} n_j c_j^o \mu_j^\star \qquad \text{and} \qquad B = \sum_{j \neq i,\ i+1} n_j c_j^{o2}.$$

We begin by comparing $\gamma(n_i \mu_i^\star + n_{i+1} \mu_{i+1}^\star)$ of (b) and $(n_i c_i^o \mu_i^\star + n_{i+1} c_{i+1}^o \mu_{i+1}^\star)$ of (a). Expanding the formula for γ and rewriting, this is equivalent to comparing

$$(n_i c_i^o + n_{i+1} c_{i+1}^o)(n_i \mu_i^\star + n_{i+1} \mu_{i+1}^\star) \quad \text{and} \quad (n_i c_i^o \mu_i^\star + n_{i+1} c_{i+1}^o \mu_{i+1}^\star)(n_i + n_{i+1}).$$

Expanding these equations and canceling like terms we have

$$n_i n_{i+1}(c_i^o \mu_{i+1}^\star + c_{i+1}^o \mu_i^\star) \quad \text{and} \quad n_i n_{i+1}(c_i^o \mu_i^\star + c_{i+1}^o \mu_{i+1}^\star).$$

Their difference may be written as

$$(c_i^o - c_{i+1}^o)(\mu_{i+1}^\star - \mu_i^\star) > 0.$$

thus the first part of (b) is greater than the first part of (a).

Next, considering the second part of equations (a) and (b) we compare

$$n_i c_i^{o2} + n_{i+1} c_{i+1}^{o2} \quad \text{and} \quad \gamma^2(n_i + n_{i+1}).$$

As before, expanding γ and rewriting, the above is equivalent to

$$(n_i + n_{i+1})(n_i c_i^{o2} + n_{i+1} c_{i+1}^{o2}) \quad \text{and} \quad (n_i c_i^o + n_{i+1} c_{i+1}^o)^2.$$

As above, we expand the expressions and cancel like terms which evaluate to

$$(n_i n_{i+1})(c_i^{o2} + c_{i+1}^{o2}) \quad \text{and} \quad 2n_i n_{i+1} c_i^o c_{i+1}^o,$$

which implies $c_i^{o2} + c_{i+1}^{o2} > 2c_i^o c_{i+1}^o$. Hence the second part of equation (b) is greater than the second part of equation (a). This implies $\mathbf{c^o}$ is not the optimal solution, which is a contradiction.

Thus $c_i^o \leq c_{i+1}^o$ as required. $\square$

An equivalence relationship is established in the following theorem between the positive nature of the lower bound $L(\mu_k - \mu_1)$ and the rejection of H_0 by statistic T. Its proof is very similar to that of Theorem 7.2.1 of Peng (2002) and thus is omitted.

Theorem 3.1.5. When $\mu \in \Omega$, $T > t_{k,\nu,\alpha}$ if and only if $L(\mu_k - \mu_1) > 0$.

When the lower bound $L(\mu_k - \mu_1)$ is positive, it signifies the mean μ_k is significantly larger than the control mean μ_1 and the size of the difference is measured. We may also implement the multiple contrast test in a procedure to find the MED by rejecting the null hypothesis if $L(\mu_k - \mu_1) > \delta$, as will be detailed in the step-wise procedure to be introduced in the next Chapter.

3.2 Optimization Theorem for Simultaneous Confidence Lower Bound

With the results from the previous section, the following Theorem establishes a necessary and sufficient condition for an optimal solution and its proof follows.

Theorem 3.2.1. Suppose $L(\mu_k - \mu_1) > 0$. Then the vector $\mathbf{c}^o \in \mathcal{K}$ is an optimal solution to (3.5) if and only if there exist non-negative integers p and q, $1 \leq p < q \leq k$, such that $\mu_p^\star < \hat{\mu} < \mu_q^\star$, $S_{1p}^2 + S_{qk}^2 < S^2 t_{k,\nu,\alpha}^2$ and $c_1^o \leq \cdots \leq c_p^o < c_{p+1}^o = \cdots = c_{q-1}^o = 0 < c_q^o \leq \cdots \leq c_k^o$, where $c_i^o = -N_{1p}^{-1} + b^{-1}(\mu_i^\star - \bar{Y}_{1p})$, $i = 1, \ldots, p$, and $c_i^o = N_{qk}^{-1} + b^{-1}(\mu_i^\star - \bar{Y}_{qk})$, $i = q, \ldots, k$, with

$$\max\{N_{1p}(\mu_p^\star - \bar{Y}_{1p}), N_{qk}(\bar{Y}_{qk} - \mu_q^\star)\} < b \leq \min\{N_{1p+1}(\mu_{p+1}^\star - \bar{Y}_{1p+1}), N_{q-1k}(\bar{Y}_{q-1k} - \mu_{q-1}^\star)\}. \tag{3.6}$$

where

$$N_{ab} = \sum_a^b n_i, \quad \bar{Y}_{ab} = \sum_a^b n_i \mu_i^\star / N_{ab}, \quad S_{ab}^2 = \sum_a^b n_i (\mu_i^\star - \bar{Y}_{ab})^2$$

and

$$b^2 = (S^2 t_{k,\nu,\alpha}^2 - S_{1p}^2 - S_{qk}^2)/(N_{1p}^{-1} + N_{qk}^{-1}). \tag{3.7}$$

When $q = p + 1$, the upper bound for b in (3.6) is $(\bar{Y}_{qk} - \bar{Y}_{1p})/(N_{1p}^{-1} + N_{qk}^{-1})$ and the lower bound remains the same.

Proof of Theorem 3.2.1.

We begin the proof by noting that Lemma 3.1.4 allows the constraints to be replaced by $\sum_{i=1}^k n_i c_i = 0$ and $\sum_{i=j}^k n_i c_i \leq 1$, $j = 2, \ldots, k$.

We have that $g(c) = \sum_{i=1}^k n_i c_i \mu_i^\star - t_{k,\nu,\alpha} S \sqrt{\sum_{i=1}^k n_i c_i^2}$ is a concave function of $c_1, c_2, \ldots, c_k$. (*Proof*: Write $\sum_{i=1}^k n_i c_i^2 = \mathbf{c}'\mathbf{N}\mathbf{c}$. From p. 48 of Liu (2001).

$(\sum_{i=1}^{k} n_i c_i^2)^{1/2}$ is convex, thus $-(\sum_{i=1}^{k} n_i c_i^2)^{1/2}$ is concave. This implies $g(\mathbf{c})$ is concave also.)

Note that $H_1(\mathbf{c}) = 1 - \sum_{i=j}^{k} n_i c_i \geq 0$, $j = 2, \ldots, k$ and $H_2(\mathbf{c}) = \sum_{i=1}^{k} n_i c_i = 0$ hence we use constraints C1 and C3 from Kuhn -Tucker: $v_j \geq 0$, $\lambda \in R$. Let

$$\phi(c, v, \lambda) \;=\; \sum_{i=1}^{k} n_i c_i \mu_i^{\star} - t_{k,\nu,\alpha} S \sqrt{\sum_{i=1}^{k} n_i c_i^2} + \sum_{j=2}^{k} v_j (1 - \sum_{r=j}^{k} n_r c_r) - \lambda \sum_{i=1}^{k} n_i c_i.$$

Furthermore, let $\dfrac{\partial \phi}{\partial \mathbf{c^o}}$ denote the partial derivatives evaluated at the point $\mathbf{c^o}, \mathbf{v^o}, \lambda^o$. By the Kuhn–Tucker equivalence theorem, $\mathbf{c^o}$ is the optimal solution if and only if

(i) $\dfrac{\partial \phi}{\partial c_i^o} = n_i \mu_i^{\star} - n_i c_i^o b - n_i \sum_{j=2}^{i} v_j^o - \lambda^o n_i \;=\; 0$, $(i = 1, \ldots, k)$,

with $b = t_{k,\nu,\alpha} S / (\sum_{i=1}^{k} n_i c_i^{o2})^{1/2}$,

(ii) $\dfrac{\partial \phi}{\partial v_j^o} = 1 - \sum_{r=j}^{k} n_r c_r^o \geq 0 \iff \sum_{r=j}^{k} n_r c_r^o \leq 1 \;(j = 2, \ldots, k)$,

$\left(\dfrac{\partial \phi}{\partial v_j^o}\right) v_j^o = 0 \iff v_j^o (1 - \sum_{r=j}^{k} n_r c_r^o) \;=\; 0$ for all $j = 2, \ldots, k$.

$\mathbf{v^o} \geq \mathbf{0}$, $\dfrac{\partial \phi}{\partial \lambda^o} = \sum_{i=1}^{k} n_i c_i = 0$.

Suppose $\mathbf{c^o}$ is the optimal solution. By Lemma 3.1.4, then

$$c_1^o \leq c_k^o \leq \cdots \leq c_k^o.$$

We may split the coefficients into those positive and negative valued c_i^o's as:

$$c_1^o \leq \cdots \leq c_p^o < c_{p+1}^o = \cdots = c_{q-1}^o = 0 < c_q^o \leq \cdots \leq c_k^o.$$

with

$$\sum_{i=1}^{p} n_i c_i^o = -1 \quad \text{and} \quad \sum_{i=q}^{k} n_i c_i = 1.$$

(*Proof.* We know that with c_i^o, $\sum n_i c_i^o \mu_i \leq \mu_k - \mu_1$, i.e. $l(\mu_k - \mu_1) = l(\sum n_i c_i^o \mu_i) > 0$. Suppose $0 < \sum_{i=q}^{k} n_i c_i^o = a < 1$, thus $\sum_{i=1}^{p} n_i c_i^o = -a$.

Let $d^o = c^o/a$. Then $\sum n_i d_i^o \mu_i \leq \sum n_i c_i^o \mu_i \leq \mu_k - \mu_1$. Therefore,

$$l(\sum n_i d_i^o \mu_i) = \frac{1}{a} l(\sum n_i c_i^o \mu_i) > l(\sum n_i c_i^o \mu_i),$$

since $a < 1$. This implies a contradiction, since c_i^o is the optimal solution, thus $d^o = c^o \iff a = 1$, satisfying the result.)

From (ii),

$$c_i^o = 0 \text{ when } \sum_{r=i}^{k} n_r c_r^o < 1 \text{ for } i = 2, \ldots, p, q+1, \ldots, k.$$

From (i),

$$c_i^o = b^{-1}(\mu_i^\star - \sum_{j=2}^{i} v_j^o - \lambda^o) \tag{3.8}$$

Adding the first p equations in (3.8) and using the result that $\sum_{i=1}^{p} n_i c_i^o = -1$, then

$$\begin{aligned} \sum_{i=1}^{p} n_i c_i^o &= b^{-1} \left(\sum_{i=1}^{p} n_i \mu_i^\star - \sum_{i=1}^{p} \sum_{j=2}^{i} n_i v_j^o - \lambda^o \sum_{i=1}^{p} n_i \right) \\ -1 &= b^{-1} \left(\sum_{i=1}^{p} n_i \mu_i^\star - \sum_{i=1}^{p} \sum_{j=2}^{i} n_i v_j^o - \lambda^o \sum_{i=1}^{p} n_i \right) \end{aligned}$$

$$
\begin{aligned}
-1 &= b^{-1}\left(\sum_{i=1}^{p} n_i\mu_i^\star - \lambda^o \sum_{i=1}^{p} n_i\right), \quad \text{since } \sum_{i=1}^{p}\sum_{j=2}^{i} n_i v_j^o = 0 \\
\lambda^o &= \frac{b}{\sum_{i=1}^{p} n_i} + \frac{\sum_{i=1}^{p} n_i\mu_i^\star}{\sum_{i=1}^{p} n_i}
\end{aligned}
$$

Therefore, the first p coefficients satisfy the equation

$$
\begin{aligned}
c_i^o &= b^{-1}(\mu_i^\star - \sum_{j=2}^{i} v_j^o - \lambda^o) \quad \text{from (3.8). so} \\
c_i^o &= b^{-1}\left(\mu_i^\star - \frac{b}{\sum_{i=1}^{p} n_i} - \frac{\sum_{i=1}^{p} n_i\mu_i^\star}{\sum_{i=1}^{p} n_i}\right) \quad i = 1,\ldots,p \text{ since } v_j^o = 0. \\
\therefore c_i^o &= b^{-1}(\mu_i^\star - \bar{Y}_{1p}) - N_{1p}^{-1} \quad i = 1,\ldots,p
\end{aligned}
$$

and

$$
\lambda^o = \bar{Y}_{1p} + bN_{1p}^{-1}.
$$

Next, let $V = \sum_{i=p+1}^{q} v_i^o$. With $\sum_{i=1}^{k} n_i c_i^o = 1$, we sum the equations from q to k as follows:

$$
\begin{aligned}
\sum_{i=1}^{k} n_i c_i^o &= b^{-1}\left(\sum_{i=q}^{k} n_i\mu_i^\star - \sum_{i=q}^{k} n_i \sum_{j=1}^{i} v_j^o - \lambda^o \sum_{i=q}^{k} n_i\right) \\
1 &= b^{-1}\left(\sum_{i=q}^{k} n_i\mu_i^\star - \sum_{i=q}^{k} n_i V - \lambda^o \sum_{i=q}^{k} n_i\right) \text{ since } \sum_{i=2}^{p} v_j^o = \sum_{i=q+1}^{k} v_j^o = 0, \\
\frac{b}{\sum_{i=q}^{k} n_i} &= \frac{\sum_{i=q}^{k} n_i\mu_i^\star}{\sum_{i=q}^{k} n_i} - V - \lambda^o \\
V &= \bar{Y}_{qk} - bN_{qk}^{-1} - \lambda^o = \sum_{j=2}^{i} v_j^o \quad \text{for } i = q,\ldots,k
\end{aligned}
$$

As before, the last q coefficients may be written as

$$\begin{aligned} c_i^o &= b^{-1}(\mu_i^\star - V - \lambda^o) \quad \text{from (3.8), thus} \\ c_i^o &= b^{-1}(\mu_i^\star - \bar{Y}_{qk}) + N_{qk}^{-1} \quad i = q, \ldots, k. \end{aligned}$$

With these equations for $\mathbf{c^o}$, $b^{-1} = (\sum_{i=1}^k n_i c_i^{o2})^{1/2}/t_{k,\nu,\alpha}S$ may be expressed as follows:

$$\begin{aligned} S^2 t_{k,\nu,\alpha}^2 b^{-2} &= \sum_{i=1}^k n_i c_i^{o2} \\ &= \sum_{i=1}^p n_i(-N_{1p}^{-1} + b^{-1}(\mu_i^\star - \bar{Y}_{1p}))^2 + \sum_{i=q}^k n_i(N_{qk}^{-1} + b^{-1}(\mu_i^\star - \bar{Y}_{qk}))^2 \\ &= \sum_{i=1}^p n_i(N_{1p}^{-2} - 2b^{-1}N_{1p}^{-1}(\mu_i^\star - \bar{Y}_{1p}) + b^{-2}(\mu_i^\star - \bar{Y}_{1p})^2) \\ &\quad + \sum_{i=q}^k n_i(N_{qk}^{-2} + 2b^{-1}N_{qk}^{-1}(\mu_i^\star - \bar{Y}_{qk}) + b^{-2}(\mu_i^\star - \bar{Y}_{qk})^2) \end{aligned}$$

However, the middle terms sum to zero, thus the equation becomes

$$S^2 t_{k,\nu,\alpha}^2 b^{-2} = N_{1p}^{-1} + N_{qk}^{-1} + b^{-2}\sum_{i=1}^p n_i(\mu^\star - \bar{Y}_{1p})^2 + b^{-2}\sum_{i=q}^k n_i(\mu_i^\star - \bar{Y}_{qk})^2.$$

It follows that

$$b^{-2} = \frac{N_{1p}^{-1} + N_{qk}^{-1}}{S^2 t_{k,\nu,\alpha}^2 - S_{1p}^2 - S_{qk}^2},$$

where $S_{1p}^2 = \sum_{i=1}^p n_i(\mu_i^\star - \bar{Y}_{1p})^2$ and $S_{qk}^2 = \sum_{i=q}^k n_i(\mu_i^\star - \bar{Y}_{qk})^2$.

To show the validity of equation (3.6), we must consider two cases: $q > p+1$

and $q = p+1$. For the case $q > p+1$, from (3.8) we have $\sum_{j=1}^{p+1} v_j^o = \sum_{j=1}^{p} v_j^o + v_{p+1}^o = 0 + v_{p+1}^o = v_{p+1}^o$ and $c_{p+1}^o = 0$ which implies $v_{p+1}^o = \mu_{p+1}^\star - \lambda^o$.

Then, $\sum_{j=1}^{p+2} v_j^o = v_{p+1}^o + v_{p+2}^o$ and $c_{p+2}^o = 0$ which implies $v_{p+2}^o = \mu_{p+2}^\star - v_{p+1}^o - \lambda^o = \mu_{p+2}^\star - \mu_{p+1}^\star + \lambda^o - \lambda^o = \mu_{p+2}^\star - \mu_{p+1}^\star \geq 0$ and so on. Similarly, $v_{q-1}^o = \mu_{q-1}^\star - \mu_{q-2}^\star \geq 0$ and $v_q^o = \bar{Y}_{qk} - \mu_{q-1}^\star - bN_{qk}^{-1} \geq 0$.

By (ii), $v_{p+1}^o \geq 0$ and $v_q^o \geq 0$ so $v_{p+1}^o = \mu_{p+1}^\star - \bar{Y}_{1p} - bN_{1p}^{-1} \geq 0$ and $v_q^o = \bar{Y}_{qk} - \mu_{q-1}^\star - bN_{qk}^{-1} \geq 0$. Therefore $b \leq (\mu_{p+1}^\star - \bar{Y}_{1p})N_{1p}$ and $b \leq (\bar{Y}_{qk} - \mu_{q-1}^\star)N_{qk}$, which implies

$$\begin{aligned} b &\leq \min\{N_{1p}(\mu_{p+1}^\star - \bar{Y}_{1p}), N_{qk}(\bar{Y}_{qk} - \mu_{q-1}^\star)\} \\ &= \min\{N_{1p+1}(\mu_{p+1}^\star - \bar{Y}_{1p+1}), N_{q-1k}(\bar{Y}_{q-1k} - \mu_{q-1}^\star)\}. \end{aligned}$$

The last equation follows since

$$\begin{aligned} N_{1p+1}(\mu_{p+1}^\star - \bar{Y}_{1p+1}) &= N_{1p}(\mu_{p+1}^\star + n_{p+1}\mu_{p+1}^\star) - \sum_{i=1}^{p+1} n_i\mu_i^\star \\ &= N_{1p}\mu_{p+1}^\star - \sum_{i=1}^{p} n_i\mu_i^\star \\ &= N_{1p}(\mu_{p+1}^\star - \bar{Y}_{1p}). \end{aligned}$$

For the case $q = p+1$, since

$$\begin{aligned} v_q^o &= \bar{Y}_{qk} - bN_{qk}^{-1} - \bar{Y}_{1p} - bN_{1p}^{-1} \\ &= \bar{Y}_{qk} - \bar{Y}_{1p} - b(N_{1p}^{-1} + N_{qk}^{-1}) \geq 0. \end{aligned}$$

then

$$b \leq (\bar{Y}_{qk} - \bar{Y}_{1p})/(N_{1p}^{-1} + N_{qk}^{-1}).$$

In addition, $c_p^o < 0$ and $c_q^o > 0$ so

$$c_p^o = -N_{1p}^{-1} + b^{-1}(\mu_p^\star - \bar{Y}_{1p}) < 0 \implies b^{-1} < N_{1p}^{-1}(\mu_p^\star - \bar{Y}_{1p})^{-1}$$

$$\implies b > N_{1p}(\mu_p^\star - \bar{Y}_{1p})$$

and

$$c_q^o = N_{qk}^{-1} + b^{-1}(\mu_q^\star - \bar{Y}_{qk}) > 0 \Longrightarrow b > N_{qk}(\bar{Y}_{qk} - \mu_q^\star).$$

Hence

$$b > \max\{N_{1p}(\mu_p^\star - \bar{Y}_{1p}), N_{qk}(\bar{Y}_{qk} - \mu_q^\star)\}.$$

Thus, equation (3.6) follows and the Theorem is proven. □

3.3 Iterative Algorithm

Of the k possible treatments, there are $\binom{k-1}{2}$ possible choices of p and q, with $1 \leq p < q \leq k$. As evident from Optimization Theorem 3.2.1, the choice of $p < q$ defines the optimal solution if the condition (3.6) holds. For given set of $\bar{Y}_1, \ldots, \bar{Y}_k$, there are no more than $k-1$ possible choices of (p, q) for the optimal solution $\mathbf{c^o}$, depending upon the confidence level $1 - \alpha$.

The following algorithm selects the optimal solution (p_0, q_0), (p_1, q_1), ..., (p_r, q_r) from confidence level $1 - p$ to the desired level $1 - \alpha$ where p is the p-value of the multiple contrast test statistic T.

(0) Set $i = 0$, $p_0 = \max\{1 \leq j < k : \mu_j^\star < \hat{\mu}\}$ and $q_0 = \min\{2 \leq j \leq k : \mu_j^\star > \hat{\mu}\}$.

(i) Let $\mathcal{B}_i = \max\{N_{1p_i}(\mu_i^\star - \bar{Y}_{1p_i}), N_{q_ik}(\bar{Y}_{q_ik} - \mu_{q_i}^\star)\}$, $t_{k,\alpha_{i+1}} = \{S_{1p_i}^2 + S_{q_ik}^2 + (N_{1p_i}^{-1} + N_{q_ik}^{-1})\mathcal{B}_{i+1}^2\}^{1/2}/S$. If $t_{k,\alpha_{i+1}} < t_{k,\alpha}$ then the optimal solution is $\mathbf{c}^{\mathbf{o}}$ with $p = p_i$, $q = q_i$. Otherwise go to (ii).

(ii) If $N_{1p_i}(\mu_{p_i}^\star - \bar{Y}_{1p_i}) > N_{q_ik}(\bar{Y}_{q_ik} - \mu_{q_i}^\star)$ then set $p_{i+1} = \max\{j : 1 \leq j < p_i, \mu_j^\star < \mu_{p_i}^\star\}$ and $q_{i+1} = q_i$. Otherwise, set $p_{i+1} = p_i$ and $q_{i+1} = \min\{j : q_i < j \leq k, \mu_j^\star > \mu_{q_i}^\star\}$. Set $i = i + 1$ and go to step (i).

The reader is asked to refer to Appendix I for an S-Plus program written to evaluate the steps of the Iterative Algorithm, as outlined above. A justification of this algorithm is now presented.

Proof of algorithm.

In a similar manner as Lee, Peng and Liu (2001), at step (0), let $p = p_0$, $q = q_0$, $\mathcal{B}_0$ be the upper bound of b, i.e. $\mathcal{B}_0 = (\bar{Y}_{qk} - \bar{Y}_{1p})/(N_{1p}^{-1} + N_{qk}^{-1})$. In

addition, let $t_{k,\nu,\alpha_0} = T$, which implies that the p-value equals α_0. Then,

$$\begin{aligned}
t^2_{k,\nu,\alpha_0} \;=\; T^2 \;&=\; \sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2/S^2 \\
\Longrightarrow\; S^2 t^2_{k,\nu,\alpha_0} \;&=\; \sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2 \\
&=\; \sum_{i=1}^{p} n_i(\mu_i^\star - \bar{Y}_{1p})^2 + N_{1p}(\bar{Y}_{1p} - \hat{\mu})^2 + \sum_{i=q}^{k} n_i(\mu_i^\star - \bar{Y}_{qk})^2 \\
&\quad + N_{qk}(\bar{Y}_{qk} - \hat{\mu})^2 \\
&=\; S^2_{1p} + S^2_{qk} + N_{1p}(\bar{Y}_{1p} - \hat{\mu})^2 + N_{qk}(\bar{Y}_{qk} - \hat{\mu})^2.
\end{aligned}$$

We note that $\hat{\mu}$ is the weighted average of $\bar{Y}_{1p}$ and $\bar{Y}_{qk}$ and thus

$$\begin{aligned}
\bar{Y}_{1p} - \hat{\mu} \;&=\; \bar{Y}_{1p} - \frac{N_{1p}\bar{Y}_{1p} + N_{qk}\bar{Y}_{qk}}{N_{1p} + N_{qk}} \\
&=\; \frac{-N_{qk}(\bar{Y}_{qk} - \bar{Y}_{1p})}{N_{1p} + N_{qk}}.
\end{aligned}$$

Similarly, $\bar{Y}_{qk} - \hat{\mu} = \dfrac{N_{1p}(\bar{Y}_{qk} - \bar{Y}_{1p})}{N_{1p} + N_{qk}}$.

Therefore,

$$\begin{aligned}
S^2 t^2_{k,\nu,\alpha_0} \;&=\; S^2_{1p} + S^2_{qk} + \left(\frac{N_{1p}N^2_{qk}}{(N_{1p} + N_{qk})^2} + \frac{N_{qk}N^2_{1p}}{(N_{1p} + N_{qk})^2} \right) (\bar{Y}_{qk} - \bar{Y}_{1p})^2 \\
&=\; S^2_{1p} + S^2_{qk} + \frac{(\bar{Y}_{qk} - \bar{Y}_{1p})^2}{N^{-1}_{1p} + N^{-1}_{qk}} \\
&=\; S^2_{1p} + S^2_{qk} + (N^{-1}_{1p} + N^{-1}_{qk})\beta_0^2.
\end{aligned}$$

Thus

$$t^2_{k,\nu,\alpha_0} \;=\; \{S^2_{1p} + S^2_{qk} + (N^{-1}_{1p} + N^{-1}_{qk})\beta^2_0\}/S^2.$$

When $t_{k,\nu,\alpha_0} \geq t_{k,\nu,\alpha} > t_{k,\nu,\alpha_1}$ then $\alpha_0 \leq \alpha < \alpha_1$ and $\beta_0 \geq b(\alpha) > \beta_1$ with

$$b(\alpha) \;=\; (t^2_{k,\nu,\alpha}S^2 - S^2_{1p} - S^2_{qk})/(N^{-1}_{1p} + N^{-1}_{qk}).$$

Then, $S^2 t_{k,\nu,\alpha_1} = S^2_{1p_0} + S^2_{q_0 k} + (N^{-1}_{1p} + N^{-1}_{qk})\beta^2_1$ with $t_{k,\nu,\alpha_0} \geq t_{k,\nu,\alpha_1}$ implies that $\beta_0 \geq \beta_1$, and b satisfies (3.6) at step (0).

Next, let $\beta_1 = N_{1p}(\mu^\star - \bar{Y}_{1p})$. Then $p_1 = p - 1$ and $q_1 = q$. We then have

$$\begin{aligned}
S^2 t^2_{k,\nu,\alpha_1} &= S^2_{1p} + S^2_{qk} + (N^{-1}_{1p} + N^{-1}_{qk})\beta^2_1 \\
&= S^2_{1p-1} + S^2_{qk} + N_{1p-1}(\bar{Y}_{1p-1} - \bar{Y}_{1p})^2 + n_p(\mu^\star_p - \bar{Y}_{1p})^2 \\
&\quad + (N^{-1}_{1p} + N^{-1}_{qk})N^2_{1p}(\mu^\star_p - \bar{Y}_{1p})^2 \\
&= S^2_{1p-1} + S^2_{qk} + (N^{-1}_{1p-1} + N^{-1}_{qk})\beta^2_1.
\end{aligned}$$

As before, $t_{k,\nu,\alpha} \leq t_{k,\nu,\alpha_1}$ implies $\alpha \geq \alpha_1$ and $\beta_1 \geq b(\alpha)$. These relationships also hold for $\beta_1 = N_{qk}(\bar{Y}_{qk} - \mu^\star_q)$ when $p_1 = p$ and $q_1 = q + 1$. By induction, the desired p_i and q_i are acquired such that (3.6) holds for the value of $b(\alpha)$ at a given level $1 - \alpha$.

3.4 Numerical Example

To illustrate the results of the previous sections, consider the binding assay data from Lee (1996), as displayed in Table 3.1. In this instance, $k = 9$ treatment levels of antiserum dilution were used and a response of % inhibition of rosettes was measured.

Table 3.1

Binding Inhibition Assay Data for Numerical Example 3.4 from Lee (1996)

$\log_{10}$ **dilution**	**Inhibition (%)**	**Dose Means** ($\bar{Y}_i$)
3.519	-12, 5	-3.5
3.114	12, 27	19.5
2.778	14, 18, 25, 36	23.25
2.399	44, 46	45
2.000	44, 45, 46	45
1.399	27, 33, 56	38.67
1.000	38, 40	39
0.699	32, 43, 50, 54	44.75
0.301	43, 47	45

The sample means and sample sizes for each dose level are calculated as:

$$\bar{\mathbf{Y}} = (-3.5, 19.5, 23.25, 45, 45, 38.67, 39, 44.75, 45)$$

$$\mathbf{n} = (2, 2, 4, 2, 3.3, 2, 4, 2)$$

We also have $S^2 = 86.477$ and $\nu = 15$. For a 95% confidence level, due to unequal sample sizes, a simulated critical value from (2.9) was obtained as $t_{9,0.05} = 2.926$ using 1,000,000 replications. Also, $\hat{\mu} = 33.875$ and the maximum likelihood estimates of the sample means are:

$$\mu^\star = (-3.5, 19.5, 23.25, 41.9, 41.9, 41.9, 41.9, 44.75, 45)$$

We next use the Iterative Algorithm of Section 3.3 to evaluate the optimal lower bound $L(\mu_9 - \mu_1)$ for the difference between treatment 9 and the control treatment. Since $\hat{\mu} = 33.875$ falls between $\mu_3^\star = 23.25$ and $\mu_4^\star = 45$ then the initial values are $p_0 = 3$, $q_0 = 4$. The algorithm proceeds as follows:

- step 0: $p_0 = 3$, $q_0 = 4$

$\beta_1 = \max\{61, 17.6\} = 61$. $t_{9,15,\alpha_1} = 4.457 \not< t_{9,.05} = 2.926$ so go to the next step.

- step 1: $p_1 = 2$, $q_1 = 4$

$\beta_2 = \max\{46, 17.6\} = 46$. $t_{9,15,\alpha_2} = 3.753 \not< t_{9,.05} = 2.926$ so continue to the next step.

- step 2: $p_2 = 1$, $q_2 = 4$

$\mathcal{B}_3 = \max\{0, 17.6\} = 17.6$, $t_{9,15,\alpha_3} = 1.542 < 2.926$, so (1.4) is the optimal choice of (p, q).

Then, from equation (2.15), the optimal simultaneous lower bound is $L(\mu_9 - \mu_1) = 26.5085$ with

$$\mathbf{c^0} = (-.5, 0, 0, .0316, .0316, .0316, .0316, .1117, .1187)$$

In the following Chapters, this algorithm is utilized in the calculation of the minimum effective dose (MED) through a step-wise testing procedure.

Chapter 4

Applications to Dose-Response Studies

In this Chapter, various aspects of biopharmaceutical research are discussed. The process of modern drug development employs numerous statistical methods as attempts are made to assess the pharmacologic activity of a compound. For a toxic substance, determination of those safe doses is desired, while for a pharmaceutical drug, interest lies in which dosage is needed for an effect. In a clinical research programme for a new chemical entity, an investigation of the dose-response relationship is more or less a mandatory component (Källén and Larsson (1999)). As mentioned by Ruberg (1995a), classical dose response stud-

ies are used in pharmacology, toxicology and clinical research. Dose response relationships and studies may also be involved with pharmacokinetics, assay validation and concentration response for *in vitro* studies.

Ruberg (1995a) notes the four fundamental questions that must be answered in studying the dose-response relationship of a drug as follows:

1. Is there any evidence of a drug effect?

2. What doses exhibit a response different from the control response?

3. What is the nature of the dose response relationship?

4. What is the optimal dose?

At the beginning of the drug development process, question 1 is of concern to decide whether to proceed with research into this compound, whereas question 4 is of primary focus to researchers nearing completion of an efficacy study.

With regard to experimental design, the most common design used in dose-response studies is the placebo-controlled, randomized, parallel-dose response study (Ruberg (1995a)). In this design, patients are randomly allocated to one of several active dose groups or placebo. The popularity of the design lies in the fact that the only difference between treatment groups is the dose of the

experimental compound, which leads to straightforward interpretation of results. A placebo group is important to the study as a significant trend in response with increasing dose in the absence of placebo is not necessarily evidence of a drug effect (Ruberg (1995a)). However, Ruberg also states situations which do not necessitate a placebo group to assess dose-response significance (Ruberg (1995a, p. 3)).

An important and practical dosing quantity in drug development studies is now defined.

4.1 The Minimum Effective Dose (MED)

4.1.1 Definition of the MED

Hsu and Berger (1999) note that in practice, dosing is determined by two quantities:

- minimum effective dose (MED)
- maximum tolerated dose (MTD)

The MED of a drug is the minimum dose such that the mean response at that dose is significantly better than the mean response of the controls. The estimated

MED is determined statistically from the observed dose-response relationship, where the response is an endpoint that measures efficacy. The estimated MTD on the other hand, is determined from observed adverse events in terms of both anticipated and unanticipated endpoints. Hsu and Berger (1999) consider the estimation of the MTD to be nonstatistical at present and thus limit their discussion to the MED problem.

If the response curve is expected to be continuous, the MED should be defined as the minimum dose such that the mean response at that dose is clinically significantly better than the mean response of the ith treatment; that is

$$MED = \min\{i : \mu_i > \mu_1 + \delta\} \tag{4.1}$$

where $\delta \geq 0$ defines a clinically significant difference. For instance, for chronic peripheral arterial occlusive disease, in terms of percent improvement in walking distance, δ has been defined to be 30% (Hsu and Berger (1999)). Ruberg (1995a) notes that a clinically important response is needed since small, statistically significant responses may be meaningless, while on the other hand, a statistically significant response is required as large clinical responses that are not clearly distinguishable from placebo response do not provide substantive evidence of drug effectiveness.

Generally, the situation $\mu_i \leq \mu_1 + \delta$ may arise for some dose level i. This occurs if the control group is active, i.e. receiving a standard drug known to be efficacious, or if the control group is negative, i.e. receiving a placebo, with $\delta = 0$. The latter case may occur if the response has an inverted U or umbrella shape, meaning that μ_i, $i = 1, \ldots, k$, first increases then decreases as i increases.

Since dose-response means increase for increasing dose, we require the assumption of monotonicity of the μ_i:

$$\mu_1 \leq \mu_2 \leq \cdots \leq \mu_k. \tag{4.2}$$

In the past two decades, the pharmaceutical industry has been perceived as conducting studies of doses that are too high to assure efficacy in their clinical trials, and as a result, lower doses than those declared statistically significantly different from a placebo may be found that are efficacious and have a wider safety margin. The Food and Drug Administration (FDA) has reported that approximately 10% of drugs (new molecular entities) that were approved in 1980-1989 have had dosage changes (mostly decreases in dose) of greater than 33% (Ruberg (1995a)). As discussed by Hsu and Berger (1999), if the $\hat{MED}$ is found to be less than the $\hat{MTD}$, then a therapeutic window of safe and effective doses is established for the range $\hat{MED}$, $\hat{MED} + 1$, ..., $\hat{MTD}$. A wide therapeutic

window of safe and effective doses is beneficial to both the manufacturer and regulatory body as it facilitates first choosing a higher dose for the prescription version and later a lower dose for the over-the-counter version of the same drug. Evidently, the need for more accurate measurements of dose efficacy has led to the determination of the MED to be a fundamental aspect of research in drug development studies.

4.1.2 Assessment of a Dose-Response Relationship

In general, analysis of the four fundamental questions in drug development, which include the minimum effective dose determination, is facilitated by two basic approaches - *hypothesis testing* which involves analysis of variance (ANOVA) followed by multiple comparisons or contrasts, or *regression modeling* followed by estimation of relevant dose parameters. These two approaches, along with advantages and disadvantages of each, are now detailed.

Hypothesis Testing (ANOVA) Method

In a regulatory environment, to find evidence of a dose-response relationship or determine the minimum effective dose, the approaches based on hypothesis testing which correspond to the confirmatory aspect in clinical trials are generally

used (Hamasaki, Isomura, Baba and Goto (2000); Ruberg (1995b)). As defined in Chapter 2, we consider a one-way ANOVA model in which a set of increasing dose levels are denoted 1, 2, ..., k where 1 corresponds to the zero or control dose level and n_i experimental units are tested at the ith dose level. We let the vector $\mu = (\mu_1, \mu_2, \ldots, \mu_k)$ denote the vector of response means, where μ_i corresponds to the response mean for the ith dose, $i = 1, \ldots, k$.

The problem of identifying the MED is formatted, in a similar manner as Dunnett and Tamhane (1998), as a sequence of hypothesis testing problems:

$$H_{0i}^{\delta} : \mu_i \leq \mu_1 + \delta \quad \text{versus} \quad H_{1i}^{\delta} : \mu_i > \mu_1 + \delta, \tag{4.3}$$

where δ denotes a clinically significant difference, for all $i = 2, \ldots, k$ in a stepwise fashion. When the null hypothesis is rejected in favour of the alternative hypothesis, there exists at least one treatment better than the control.

Stepwise procedures may be divided into two general types: step-down and step-up. A step-down procedure begins by testing the overall intersection hypothesis and then steps down through the hierarchy of implied hypotheses. If any hypothesis is not rejected, then all of its implied hypotheses are retained without further tests; thus a hypothesis is tested if and only if all of its implying hypotheses are rejected. On the other hand, a step-up procedure begins by test-

ing all minimal hypotheses and then steps up through the hierarchy of implying hypotheses. If any hypothesis is rejected, then all of its implying hypotheses are rejected without further tests; thus a hypothesis is tested if and only if all of its implied hypotheses are retained (Hochberg and Tamhane (1987)). Hochberg and Tamhane (1987) also note that step-down procedures are generally more powerful than the corresponding step-up procedures. Furthermore, there is a more solid theoretical foundation for the use of step-down procedures, as will now be discussed.

In determining the MED, hypothesis testing procedures should strongly control what is known as the family-wise error rate (FWE) at level of significance α. The family-wise error rate is defined as the probability of making any erroneous significance conclusion in the given family of inferences at level α (Hochberg and Tamhane (1987)). Marcus, Peritz and Gabriel (1976) derived the closure method, which constructs step-down testing procedures and leads to closed testing procedures. This method tests all possible intersections of null hypotheses, each at level α, rejecting a resulting hypothesis only if it and all other resulting hypotheses implying it are rejected. The authors also state the following theorem:

Theorem 4.1.1 (Marcus, Peritz and Gabriel (1976)). The aforementioned closed testing procedure strongly controls the Type I FWE at level α.

- *Step-down Testing Procedure*

With reference to the MED problem, suppose that H_{oj} will be rejected when the one-sided lower bound $L(\mu_j - \mu_1) > \delta$ (or the test statistic $T \geq t_{j,\nu,\alpha}$ where $t_{j,\nu,\alpha}$ is the critical value, by Theorem 3.1.5 when $\delta = 0$). Under a one-way model, the MED testing procedure has the form (modified from Liu (2001) and Hsu and Berger (1999)):

Step 1:

If $L(\mu_k - \mu_1) > \delta$,

then assert $\mu_k > \mu_1 + \delta$ and go to step 2;

else assert that there is no dose level which is significantly better than the control dose level and stop.

Step 2:

If $L(\mu_{k-1} - \mu_1) > \delta$,

then assert $\mu_{k-1} > \mu_1 + \delta$ and go to step 3;

else assert $MED = k$ and stop.

$\vdots$

Step $k-1$:

If $L(\mu_2 - \mu_1) > \delta$,

then assert $\mu_2 > \mu_1 + \delta$ and $MED = 2$;

else assert $MED = 3$ and stop.

To better understand how this stepwise method operates, let step M $(1 < M \leq k-1)$ be the step at which the stepwise method stops. If $M > 1$, then the stepwise method declares doses $k - M + 2, \ldots, k$ to be efficacious and will provide lower confidence bounds for $\mu_i - \mu_1$ when $i = k-M+2, \ldots, k$. If $M < k$, then the stepwise method fails to declare doses $2, \ldots, k-M+1$ to be efficacious and will not yield a lower confidence bound for $\mu_i - \mu_1$ when $i = 2, \ldots, k-M+1$. If $M = k$, then the stepwise method provides a lower bound for the efficacy of every dose.

Since the sequence of hypotheses is hierarchical in nature, the approach of stepping down through each hypothesis in the sequence beginning with dose k produces a closed testing procedure. Thus, by Theorem 4.1.1, the FWE is strongly controlled at nominal level α.

Another aspect of hypothesis testing that is of statistical interest is the choice between one-sided and two-sided tests. Ruberg (1995a) summarizes the opinions

of most statisticians and states that the issue is relevant for all phases of drug development. He notes that with one exception, researchers favour the use of one-sided hypotheses for comparing experimental therapies versus placebo. As a general rule, the authors who favour one-tailed hypotheses argue that the most important error to control is Pr(drug approval | drug is not effective). Since a drug could never be approved when it is less effective than placebo, such an error will not occur in the approval process, and a one-sided alternative hypothesis should be used. Furthermore, from a statistical perspective, a one-sided test is more powerful, as it more readily detects the positive drug effect.

Interestingly, the FDA prefers two-tailed hypotheses in this setting, hence clinical trials for new drugs are actually using a 0.025 significance level for hypothesis tests. As Ruberg (1995a) remarks, the conservatism implies that by switching to one-tailed hypotheses at $\alpha = 0.05$, approximately 25% fewer patients could be utilized. Thus, one-tailed hypotheses should be the default approach to designing dose-response studies, which will reduce sample size requirements and expedite the drug development and approval process.

In summary, as noted by Ruberg (1995b), the ANOVA approach is computationally simple, understandable and easily communicated to scientific colleagues. Furthermore, no specification of the functional form of the dose-response rela-

tionship is required. When these procedures are used as part of the primary analysis of dose-response data, they should be applied regardless of the significance of the overall F-test for drug effect, however there are differing views on this matter. Hochberg and Tamhane (1987, p. 108) summarize the opinions and note that in 1977, Scheffé suggested that in genuine multiple comparison problems, any inferences of interest based on the S-intervals should be pursued regardless of the outcome of the preliminary F-test.

By contrast, Hamasaki et al. (2000) state that while tests are useful for detecting evidence against the null hypothesis in the direction of a positive trend, and have concise interpretations, relatively little insight is provided into the shape of the dose-response relationship. Ruberg (1995b) states that one of the disadvantages of this approach is that inference is made only at the studied doses, and in particular, the MED can only be declared at one of the studied doses. In fact, the minimum effective dose found by the method of hypothesis testing has been termed the minimum detectable dose (MDD) (Liu, (2001)). We next discuss regression modeling as a method of assessing dose response.

Regression Models

It is the belief of some authors that proper analysis of a dose response study involves estimating the dose-response curve (Källén and Larsson (1999)). In terms of the four fundamental questions in dose-response analysis, if enough distinct dose groups are studied, regression analysis can be used to answer Question 3 - "What is the nature of the dose-response relationship?" (Ruberg 1995b). It is assumed that the dose-response relationship is $y = f(x) + \epsilon$ where $f(x)$ is monotonic and smooth with $E(\epsilon) = 0$, $\text{var}(\epsilon) = \sigma^2$ where ϵ is normally distributed. Among the most popular dose-response curve is sigmoidal, from which there are a variety of equations to describe such a dose-response relationship. Of particular interest is the four parameter logistic function given as follows (Ruberg 1995b):

$$f(x) = \frac{A - D}{[1 + (x/C)^B]} + D. \tag{4.4}$$

With such assumptions, standard nonlinear regression techniques may be used to estimate parameters. In the four-parameter logistic model, for example, C is equal to the ED_{50}, which is the dose producing a 50% response. With the parameter estimates, it is possible to characterize whatever is desirable about the dose-response relationship, including the MED or the minimum dose with

maximal effect. If the quantities are well defined, inverse regression may be used to estimate the doses producing such effects (Ruberg (1995b)).

Recently, (Källén and Larsson (1999)) also assumed the drug under study to have a monotone dose response curve with increasing dose (D). With the assumption of a limit to the amount of effect obtained by the drug through increasing the dose, the dose-response curve should start ($D = 0$) at some level E_0 and end asymptotically ($D \to \infty$) at another level denoted $E_0 + E_{\max}$. The authors also obtain the well-known sigmoid $E_{\max}$ model, given by

$$E = E_0 + \frac{E_{\max}}{1 + (ED_{50}/D)^b}.$$

The authors note that in the description of this dose-response model, there are four parameters that are interpretable in clinical terms. The parameter b is a sensitivity measure for the response variable with respect to relative increases in dose and is sometimes called the Hill parameter of the dose-response curve. If a value for $E_{\min}$, the smallest clinically meaningful effect is obtained, then the dose which produces this effect can be measured. For this estimate to be reliable, $E_{\min}$ should be in the approximately log-linear part of the dose-response curve. Furthermore, when an active control exists, estimation of the clinician's MED involves defining $E_{\min}$ for the new drug to be the same as the effect of the active

control which clinical experience has deemed the minimal effective dose. (Källén and Larsson (1999)).

Ruberg (1995b) identifies two approaches for estimating the MED from a continuous function. The first, by Davidian, Carroll and Smith (1988), was developed for assay detection limits. The relevant calculations involve the following steps: Referring to equation (4.4), let $\bar{Y}_M$ be the mean of a sample of N subjects given the MED and s_M^2 and s_A^2 be the estimated variances of $\bar{Y}_M$ and $\hat{A}$ respectively, where A is the value of the response at the lowest available dose. After constructing the appropriate t statistic defined by

$$t = \frac{\bar{Y}_M - \hat{A}}{[(s_M^2)/N + s_A^2]^{\frac{1}{2}}},$$

then, with $P(t > t_c) = \alpha$, by placing t_c in the probability equation, and solving for $\bar{Y}_M$ one obtains

$$\bar{Y}_M = \hat{A} + t_c[(s_M^2)/N + s_A^2]^{1/2}.$$

The calculated MED is then found to be $f^{-1}(\bar{Y}_M) = \hat{C}'[(\hat{A} - \hat{D})/(\bar{Y}_M - \hat{D}) - 1]^{1/\hat{B}}$.

Secondly, the segmented parabolic model is also described, which is useful when considering the low end of the response curve (Ruberg (1995b)). Essentially, one assumes there is a horizontal dose-response relationship over a low dose range, which then becomes a second-degree polynomial beyond the thresh-

old dose, denoted X_0. The functional form may be written as:

$$f(x) = \begin{cases} r_0 & x \leq X_0 \\ a + bx + cx^2 & x > X_0 \end{cases}$$

For the model to be smooth and continuous at X_0, $X_0 = -b/2c$. The dose X_0 may be interpreted as the MED, i.e. the lowest dose at which the effect of the compound under study produces a response that is different than a zero dose. More often, X_0 is referred to as the threshold rather than the MED, since the phrase "minimum effective dose" usually is thought to have a biologically or clinically meaningful effect. Ruberg (1995b) advises that when interpreting the model, the MED is the smallest dose greater than X_0 that produces an expected response greater than the clinically meaningful response.

Källén and Larsson (1999) detail marginal models for log linear and non-linear relationships. The former uses the analysis of variance model to reduce the data to treatment means, from which a log-linear model is fit to the data. In this instance, the log-linear model is an approximation of the dose-response relationship, fitted to the estimated means for the purpose of estimating the equivalent dose. The non-linear case also estimates the mean vector from analysis of variance, however it fits the sigmoidal function to the estimated mean vector. Furthermore, when there is information on individual dose-response curves, i.e.

each patient has been given more than one dose of the drug, hierarchical mixed models may be used to analyze the dose-response relationship.

Hamasaki et. al (2000) also consider regression to estimate the curve of the dose-response relationship. The authors introduce a model-based approach using a data-adaptive distribution to estimate the dose-response curve for a categorical response. They argue that this type of response is encountered more often in clinical trials than the continuous response. In addition, the descriptive empirical cumulative function determines the shape and location of the dose-response curve. The distribution function is not assumption-dependent, is a parametric distributional specification, may describe data even when random sampling is not involved and may be used directly and valuably in connection with censored samples.

In comparison with the hypothesis testing approach, regression modeling has many advantageous features. Appropriate models can describe the nature of the association, provide parameters for describing the strength of the relationship, provide predicted probabilities for the response categories at any dose, help determine the optimal dose, provide diagnostic tools for checking model assumptions and help interpret the phenomenon properly and avoid misunderstandings (Hamasaki et al. (2000)).

However, model-based approaches require many assumptions, compared to the relatively few required for test-based procedures. Furthermore, specification of the functional form of the dose-response relationship is considered difficult, and in some cases there may be several candidate models that appear to fit the data equally well but have drastically different properties at the extremes of the dose-response curve. Regression analysis may also be complicated by fitting several candidate models, by estimating weights to be used in the analysis and by difficulty in convergence of the nonlinear model to a unique solution. For most models, confidence intervals for the MED are quite broad, even when the model fits well (Hamasaki et al, (2000) and Ruberg (1995b)).

In light of the aforementioned disadvantages of using the regression model to identify the MED, and the use of the hypothesis testing approach in practice of modern clinical trials, the ANOVA approach using a step-down testing procedure will by discussed in the remainder of this paper. The next section details previous methods implemented by various authors to determine the MED. Such procedures available in the literature use test statistics or confidence bounds in the previously defined step-down testing method.

4.2 Determination of the MED by Hypothesis Testing

Previous authors have considered both likelihood ratio tests and multiple comparison tests in a stepwise testing procedure. However, few authors have considered tests in which the clinically significant difference, δ, is nonzero. A disadvantage of hypothesis testing procedures is the traditional use of point-zero (i.e. $\delta = 0$) null hypotheses for comparing a dose group with a placebo which ignores the possible difference between statistical significance and clinical importance (Hothorn and Bretz (2000)). The relevance of this omission lies in the fact that in clinical trials, statisticians and clinicians may not have an identical interpretation for the MED; the statistician may use it to mean the smallest dose that is effective from a particular study whereas the clinician may use it to mean the smallest dose that has a clinically meaningful effect (Källén and Larsson (1999)). It is hoped that the inclusion of the clinically significant difference in this paper will provide results with more meaningful conclusions and lead to the development of drugs with fewer revisions of recommended dose after promotion. The most common approaches to MED determination are now presented along with a brief discussion of each procedure.

(i) Dose-Response (DR) Method (Hsu and Berger (1999))

The DR method uses stepwise confidence intervals based on the pairwise t test statistic. The authors utilized a fundamentally different confidence set-based justification by partitioning the parameter space naturally and using the principle that exactly one member of the partition contains the true parameter. The test statistic is as follows:

$$T_j = \frac{\bar{Y}_j - \bar{Y}_1 - \delta}{s(n_j^{-1} + n_1^{-1})^{1/2}} \tag{4.5}$$

Hsu and Berger(1999) also note that the multiplicity adjustment is not needed for testing this problem as the step-down procedure is a closed testing procedure and the appropriate hypotheses are nested in sequence, beginning with the most restrictive. Thus, the critical value, $t_{\alpha,\nu}$, is the upper 100α percentile of the Student's t distribution with $\nu = \sum_{i=1}^{k} n_i - k$ degrees of freedom.

Another closed test, denoted the MPGN method, combines ideas of the multiple tests from Marcus, Peritz and Gabriel (1976) and the ranking and selection methods from Naik (1975). For this case, Hsu and Berger(1999) let $[2], \ldots, [k]$ be random indices such that T_j above may now be arranged as $T_{[2]} \leq \cdots \leq T_{[k]}$. Thus, each step in the step-down procedure is replaced by $T_{[i]} \geq d_{[i]}$, where $d_{[i]}$ is the critical value of Dunnett's (1955) (union-intersection) test.

Hsu and Berger (1999) state their simulation results indicate that the DR method tends to infer an MED that is closer to the true MED than the MPGN method or the method of Dunnett (1955). However, the DR method does not incorporate the prior knowledge that the mean responses are monotone increasing, thus may not be the ideal procedure. Two methods which utilize the monotonicity of the dose-response means are Bartholomew's Likelihood ratio test and Williams' procedures. Note that these procedures were developed for $\delta = 0$.

(ii) Williams' Procedures (Williams (1971, 1972))

Williams proposed two test statistics which take advantage of the power of isotonic regression. Williams' test statistics use isotonic estimates of the $\bar{Y}_i$'s $(i = 2, \ldots, k)$ as opposed to the original sample means. The test statistic proposed by Williams (1971) is

$$W_j^{(1)} = \frac{\mu_j^\star - \bar{Y}_1}{s(n_j^{-1} + n_1^{-1})^{1/2}},$$

Furthermore, Williams (1971, 1977) defined another test statistic which implements the isotonic estimates over all k doses as follows:

$$W_j^{(2)} = \frac{\hat{\mu}_j^\star - \hat{\mu}_1^\star}{s(n_j^{-1} + n_1^{-1})^{1/2}},$$

where

$$\mu_i^\star = \max_{2\le s\le i}\min_{i\le t\le k}\sum_{j=s}^{t} Y_j/(s-t), \qquad i=2,\ldots,k$$

$$\hat{\mu}_i^\star = \max_{1\le s\le i}\min_{i\le t\le k}\sum_{j=s}^{t} Y_j/(s-t), \qquad i=1,\ldots,k$$

For statistic $W_j^{(1)}$, Williams (1971) tabulated the upper α critical values for equal sample sizes for selected values of j, α and ν. An empirical formula to extend these to the unequal sample size case is given in Williams (1972). For statistic $W_j^{(2)}$, if σ is known, Marcus (1976) tabulated exact upper 5% and 1% quantiles for $k = 2,\ldots,5$ and estimated upper 5% and 1% quantiles for $k = 6,\ldots,11$. Williams (1977) also tabulated approximated critical values of $W_j^{(2)}$ with different degrees of freedom. As noted by Liu (2001), the approximate critical values given by Williams (1977) will result in a slight decrease in the true size and power of the test, thus use of the values given by Marcus (1976) is recommended.

(iii) Bartholomew LRT for simple order alternative (Robertson et. al. (1988))

As explained in Section 2.1.2, the likelihood ratio test with ordered alternatives is

$$S_{01} = \frac{\sum_{i=1}^{k} n_i(\mu_i^\star - \hat{\mu})^2}{\sum_{i=1}^{k} n_i(\bar{Y}_i - \hat{\mu})^2/\nu + S^2}, \tag{4.6}$$

where $\hat{\mu} = \sum_{i=1}^{k} n_i\bar{Y}_i / \sum_{i=1}^{k} n_i$. Critical values of S_{01} are tabulated in Robertson, Wright and Dykstra (1988).

In general, due to the difficulty of including a non-zero clinically significant difference (δ) value into the expressions for $W_j^{(2)}$ and S_{01}, these procedures will not be employed in numerical examples or simulation exercises.

The following procedures may be classified according to the type of contrast used in the step-down testing procedure. We recall from Section 2.3 that the $1-\alpha$ one–sided confidence bound for any contrast $\sum_{i=1}^{k} n_i c_i \mu_i$ is given by

$$l(\sum_{i=1}^{k} n_i c_i \mu_i) = \sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S (\sum_{i=1}^{k} n_i c_i^2)^{1/2} \tag{4.7}$$

For the next three procedures considered, it is noted that since the family of hypothesis under study is a closed family, a closed step-down procedure that controls the FWE and does not require ordering of the t-statistics uses an ordinary α-level t-test, i.e. the critical value is the upper α critical point of the

Student's t-distribution with ν degrees of freedom (Tamhane, Hochberg and Dunnett (1996)). Thus, the critical value is identical to the one used in the DR Procedure (Hsu and Berger (1999)).

(iv) Linear Contrast Procedure (Rom, Costello and Connell (1994))

At the jth step of the testing procedure, the general form of the linear contrasts is given by

$$c_{ij} = \begin{cases} -j+1 & i = 1: \\ c_{i-1,j} + 2 & i = 2, \ldots, j; \\ 0 & i = j+1, \ldots, k \end{cases}$$

(v) Helmert contrasts (Ruberg (1989))

The jth Helmert contrast compares the jth dose response mean with the average of all the lower dose response means, including the control, and is given by

$$c_{ij} = \begin{cases} -1 & i = 1, 2, \ldots, j-1; \\ j-1 & i = j: \\ 0 & i = j+1, \ldots, k \end{cases}$$

(vi) Reverse Helmert contrasts (Tamhane, Hochberg and Dunnett (1996))

The jth reverse Helmert contrast compares the average of the first j dose

response means with the control mean as

$$c_{ij} = \begin{cases} -j+1 & i = 1; \\ 1 & i = 2, \ldots, j; \\ 0 & i = j+1 \ldots, k \end{cases}$$

(vii) Multiple Contrast Testing Procedure

Thus, as S_{01} incorporates the monotonicity assumption of the response means, it is a more powerful test statistic for the problem when $\delta = 0$. However, due to its complex structure, the clinically significant difference δ cannot be implemented into this procedure, and as a result, construction of a confidence interval under the assumption of ordered restrictions which recognizes various values of δ is of interest.

At any step of the step-down testing procedure, when the null hypothesis is rejected in favour of the alternative hypothesis, there exists at least one treatment better than the control by the amount δ. Since $\mu_k - \mu_1$ is the largest difference between any treatment mean and the control mean, the confidence lower bound for $\mu_k - \mu_1$ is bounded below by that for any $\mu_i - \mu_1$ $(i = 2, \ldots, k)$ or their non-negative linear combinations. If this maximized confidence lower bound for $\mu_k - \mu_1$ is at least δ, then μ_k is significantly larger than $\mu_1 + \delta$ and the null hypothesis is rejected. As defined in Section 2.3, the optimal simultaneous

confidence lower bound for $\mu_k - \mu_1$ is

$$L(\mu_k - \mu_1) = \max_{\mathbf{c}\in\mathcal{K}}\{\sum_{i=1}^{k} n_i c_i \bar{Y}_i - t_{k,\nu,\alpha} S(\sum_{i=1}^{k} n_i c_i^2)^{1/2}\}.$$

It was also shown in Section 2.3 that the corresponding multiple contrast test statistic, T^2 is asymptotically equivalent to the LRT statistic S_{01}.

For all values of δ, we now implement the optimal simultaneous lower bound, as outlined in Chapter 3, which will incorporate the monotonicity of the dose–response means. The Iterative Algorithm defined in Section 3.3 may be used to calculate $L(\mu_j - \mu_1)$ at the jth step of the testing procedure. As the optimal procedure maximizes the difference between treatment groups, we expect the procedure to have a higher power in detecting the MED compared to the methods discussed previously.

4.3 Numerical Example

In this section, we apply some of the procedures outlined in the preceding section to find the minimum effective dose (MED) for a set of artificial data with $k = 7$ dose levels, including a control level. We consider each dose level to have a common sample size of $n_i = 6$ independent observations with mean squared error $S^2 = 52.25$ and degrees of freedom $\nu = 35$. The sample response means

and isotonic regression are given as follows:

$$\begin{aligned} \bar{\mathbf{Y}} &= (0, -1, 1, 10.8, 19, 20) \\ \mu^{\star} &= (-0.5, -0.5, 1, 9.9, 19, 20) \end{aligned}$$

To find the MED, we test hypothesis (4.3) beginning with $k = 7$ and consider the clinically significant difference, $\delta = 2.5$. Then, Table 4.1 displays the 95% lower confidence limits of $\mu_j - \mu_1$, $j = 7, 6, 5, 4, 3, 2$, as calculated by the methods of the previous section. The values given in parentheses indicate the actual confidence lower bound found by each method. We note that as in Hsu and Berger (1999), for compatibility, the inference given by Williams' $W(1)$ procedure is presented in terms of its associated confidence bounds.

Table 4.1

95% Simultaneous Confidence Lower Bounds on $\mu_j - \mu_1$ for Example 4.3

Groups Compared	DR Method	Williams	Linear Trend	Helmert	Reverse Helmert	Multiple Contrast
7-1	2.5 (12.95)	2.5 (12.40)	2.5 (13.44)	2.5 (8.45)	2.5 (4.12)	2.5 (12.88)
6-1	2.5 (11.95)	2.5 (11.45)	2.5 (9.92)	2.5 (9.94)	- (1.94)	2.5 (10.86)
5-1	- (0.95)	- (1.49)	2.5 (3.75)	- (-0.07)	- (-1.07)	2.5 (2.83)
4-1	- (2.95)	- (1.53)	- (2.43)	- (4.24)	- (-2.42)	2.5 (2.63)
3-1	- (-6.05)	- (-6.39)	- (-6.05)	- (-4.60)	- (-6.10)	- (-5.79)
2-1	- (-8.05)	- (-8.05)	- (-8.05)	- (-8.05)	- (-8.05)	- (-8.05)
MED	6	6	5	6	7	4

It is evident from Table 4.1 that the step-down testing procedure used with the Dose-Response, Williams (1971) and Helmert Contrast procedures finds the minimum effective dose to be dose # 6, with $\delta = 2.5$. The Reverse Helmert finds dose # 7 to be the MED, while the Linear Contrast procedure gives the MED

as dose # 5. However, the Multiple Contrast procedure performs better than all other procedures as it identifies dose # 4 to be the minimum effective dose. It is of interest, then, to determine through simulation studies the situations for which the innovative Multiple Contrast Test is more powerful and to what extent the value of δ affects the power of each procedure. Chapter 5 summarizes the simulation study performed to address these relevant issues.

4.4 Further Approaches to MED Characterization

Notwithstanding the extensive number of procedures used to determine the minimum effective dose, it remains a fertile area for research. For instance, Williams' (1971, 1972) procedures have been modified by various authors in recent literature. Shirley (1977) and Williams (1986) extended Williams' test for identifying the MED using isotonic estimators of the Kruskal-Wallis (1952) average ranks under the assumption of a monotone dose-response relationship. Bretz and Hothorn (2000) generalized Williams' test to the unbalanced case by implementing two multiple contrast tests for estimating the minimal toxic dose (MTD), which is analogous to the MED. The authors conducted a simulation

study which concluded that Williams' test is lacking in power for concave profiles in relation to both multiple contrasts, but performs well for convex dose-response shapes. Tsai (2000) proposed a test procedure similar to Williams' for testing dose-response which is based on the robust estimate of the average response. The author uses an M-estimator and a trimmed estimator in a simulation study, from which the robust method is found to be more resistant to outliers and more powerful than the Williams' procedure when the data distribution deviates from normality. However, the author notes that even though the data collected from clinical trials usually have some deviation from normality, using a higher percentage of trimming in the trimmed estimator may not be encouraged by regulatory agencies.

Further, as noted by Ruberg (1995b), while the analysis strategies use continuous data and normal theory methodologies, the concepts presented may be generalized to categorical data as well as nonparametric analysis strategies. Chuang-Stein and Agresti (1997) provide a tutorial which reviews methods for detecting a monotone dose-response relationship using discrete levels of a dose and an ordered categorical variable. These authors also discuss the modeling approach for identifying the MED and other aspects of dose response studies including stratified data, sample size considerations and small-sample methods.

Moreover, while the present discussion has focused on MED determination under the assumption of equal variance across dose levels, a recent paper by Tao, Shi, Guo and Gao (2002) extended the MED identification problem to the case of heteroskedasticity and showed the procedure properly controlled the family-wise error rate.

Secondly, the development of non-parametric procedures for characterizing the MED has been widely studied. In addition to the aforementioned procedures proposed by Shirley (1977) and Williams (1986), the Jonckheere (1954)-Terpstra (1952) test a well-known nonparametric test used for the ordered alternative problem. The Jonckheere-Terpstra test statistic sums various Mann-Whitney (1947) statistics computed between samples from the ith and jth populations under study. Neuhäuser, Liu and Hothorn (1998) modified Jonckheere-Terpstra's test by instead summing Mann-Whitney statistics computed between the two pooled groups 0 through i and $i+1$ through k. The authors also introduced a maximum test which has as a new test statistic the maximum of the modified statistic, powerful for linear and convex shapes, and the Fligner and Wolfe (1982) statistic which is powerful for concave and umbrella shapes. Simulation studies demonstrated that this maximum test should be used for concave or umbrella-shaped responses since in these cases the Jonckheere test as well as the

modified test are not very powerful. In addition, Chen (1999) proposed a multiple test based on the Mann-Whitney statistics incorporated into the step-down testing procedure. Chen found that the power performance of this test is at least as competitive to that of the isotonic regression-based methods for an ordered dose-response relationship and is more powerful than the Chen-Wolfe (1993) multiple test procedure for an umbrella ordering. Neuhäuser (2002) implemented a modified version of the recently-introduced Baumgartner-Weiß-Schindler (1998) statistic in a closed testing procedure. A simulation study showed this method to be more powerful than the Mann-Whitney or Wilcoxon test for identifying the MED. Finally, Chen and Jan (2002) proposed a non-parametric step-down closed testing procedure which extended the procedure of Chen (1999) for the randomized block design with repeated observations.

In the next Chapter, we revisit the parametric methods for MED identification as outlined in Section 4.2 and conduct a simulation study to compare the powers of the relevant procedures.

Chapter 5

Simulation Study

As previously noted, there exist many diversified procedures in the literature designed to detect the minimum effective dose. Some procedures involve specific contrasts of the response means while others involve multiple comparisons of the sample means with or without the assumption of monotonicity. A simulation study is conducted to evaluate the performance of the procedures and make recommendations as to which method is more powerful for a particular dose-response curve.

Simulation studies have been considered by various authors in order to identify effective MED detection methods. Ruberg (1995) used simulation to study the behaviour of various contrast procedures together with Williams (1971) and

Dunnett's test. Tamhane, Hochberg and Dunnett (1996) and Dunnett and Tamhane (1998) performed a more in-depth study and the latter paper found the procedure based on the likelihood ratio test (LRT) to be of superior power. Hsu and Berger (1999) evaluated their Dose-Response (DR) method under monotone and nonmonotone dose-response relationships, while Liu (2001) performed a study which computed the powers of the above tests for small sample sizes and concluded that the LRT was the most powerful test for this problem. In addition, Liu (2001) determined that the methods based on Helmert and Reverse Helmert contrasts performed much lower than other methods, and as expected, Williams' (1971, 1972) procedures were more powerful than the DR method. However, all of the aforementioned studies assumed the clinically significant difference δ to be zero, for which misunderstandings may arise between the statistician and the clinician as outlined in Section 4.2. Hence, the simulation study undertaken in this paper will compare the powers of the most common procedures to those of the innovative Multiple Contrast Method for numerous cases and differing values of the clinical difference δ.

5.1 Design of the Simulation Study

A simulation is undertaken to study the behaviour of the step-down testing procedure based on the following methods of Section 4.2: Contrast based methods including those determined by linear contrasts (denoted LC), Helmert contrasts (HC), Reverse Helmert contrasts (RH) and by General comparison methods such as the Dose-Response (DR) method, Williams (1971) procedure (W(1)) and the newly defined Multiple Contrast (MC) method.

For the study, the number of dose groups (including the control) was fixed at either $k = 6$, $k = 5$ or $k = 4$. For each value of k, with $\alpha = 0.05$ a common sample size of n was assumed for each group. In a similar manner as Tamhane, Hochberg and Dunnett (1996), μ_1 was fixed at 0, the standard error of the means, $\sigma/\sqrt{n}$ is set equal to 1 and the case of infinite degrees of freedom is considered. The study examines both linear and step functional form for the dose-response means. For a given type of response, the value of the largest mean, μ_k, is fixed at 5 and we will consider values of the clinically significant difference, δ to be 1.0, 1.5, 2.0 and 2.5. The simulation was replicated $10,000$ times and the probabilities of detecting a significant difference between the particular dose level and the control dose from the step-down testing procedure were determined.

5.2 Results of the Simulation Study

For each value of k, the probabilities of identifying the true MED from the step-down testing procedures are tabulated in Table 5.1 (for $k = 6$), Table 5.2 (for $k = 5$) and Table 5.3 (for $k = 4$). We first evaluate the estimated family-wise error (FWE) rate for each procedure. We adopt the definition of Tamhane, Hochberg and Dunnett (1996) and Dunnett and Tamhane (1998) and estimate the FWE as the proportion of replications (out of $N = 10,000$) corresponding to noneffective doses. With a linear and step dose-response curve, the estimated FWE rates are shown in Tables 5.1, 5.2 and 5.3. For a step dose-response function, the maximum value of FWE occurs when $\delta = 0$, since as δ increases for each case, the estimated FWE decreases accordingly. Furthermore, cases for which the true MED = 2 do not result in a type I error, thus the estimated FWE=.0000 is not tabulated. Finally, we remark that for all methods, the estimated FWE is less than $.05 \pm 1.96\sqrt{(.05)(.95)/10,000} = .0543$ and hence the FWE is accurately controlled at $\alpha = 0.05$.

With regard to power calculations, we note the existence of similar trends among the the three tables. For a linear response curve, the probabilities of detecting the MED are very low for all values of δ and k. In addition, we note

that the Helmert and Reverse Helmert contrasts perform very well for the cases of one effective dose and all doses effective respectively, however perform poorly for other cases. For example, in Table 5.1, for case $(0, 0, 0, 0, 0.5)$, the Helmert contrast has the highest probabilities at .7406, .8637, .9379 and .9704 whereas the Reverse Helmert has the lowest probabilities at .0010, .0042, .0162 and .0454, much lower than the other methods. Conversely, for the case $(0, 5, 5, 5, 5, 5)$, the Helmert contrast probabilities are near zero, while the Reverse Helmert contrast probabilities are the largest at .4912, .6398, .7752 and .8662. Since the Helmert and Reverse Helmert contrasts perform poorly for those cases for which they were not designed, these two methods are not employed in further analysis and are not recommended in a general MED identification procedure for which the shape of the response curve is unknown.

More specifically, an examination of the behaviour of the Multiple Contrast (MC) procedure is of interest. From Table 5.1, it is noted that the MC procedure is best in 16/24 cases and worst in only 3/24 cases. Furthermore, the maximum gain in probability over the DR, Williams and Linear Contrast procedures are .2841, .1592 and .5056 respectively, while the maximum loss in probability over these procedures is .0250, .1212 and .0309 respectively. For $k = 5$, the MC procedure is best in 13/20 cases and worst in 3/20 cases. The maximum improvements

in probability are .2392, .1432 and .3785 whereas the maximum losses in probability are .0535, .0467 and .0334 for the DR, Williams and Linear Contrast methods respectively. Finally, when $k = 4$, the Multiple Contrast procedure is best in 14/16 cases and worst in 2/16 cases. The maximum gains and maximum losses in probability for the DR, Williams and Linear Contrast procedures are .1592, .1048, .2790 and .0524, .0428, .0572 respectively. Hence, although the MC method does not demonstrate a uniform gain in probability of detecting the true MED, the improvements are substantial over the losses in probability.

As dose-response researchers are also interested in obtaining a therapeutic window of effective doses, we investigate the probabilities of correctly identifying at least one effective dose for the three values of k considered. The results are displayed in Table 5.4 ($k = 6$), Table 5.5 ($k = 5$) and Table 5.6 ($k = 4$). As in the preceding tables, since the Helmert and Reverse Helmert contrasts have significantly lower probabilities in most instances, they are omitted from the analysis. Thus, for $k = 6$, the MC procedure is best in 15/24 cases and worst in 3/24 cases. As before, the maximum gain in probability over the DR, Williams and Linear Contrast procedures are .2842, .2132 and .5168 while the maximum loss in probability to the procedures is .1236, .1048 and .0514 respectively. Similarly for $k = 5$, the MC procedure is best in 12/20 cases and worst in only 4/20

cases. The maximum increases in probability are .2259, .1751 and .3793 whereas the maximum decreases in probability are .1330, .1088 and .0753 for the DR, Williams and Linear Contrast procedures respectively. For $k = 4$, the MC procedure is best in 7/16 cases and worst in 4/16 cases. The maximum gains and losses in probability over the DR, Williams and Linear Contrast methods are .1739, .1259, .2793 and .1391, .0887, .0484 respectively. Thus, as in the previous situation, the MC procedure demonstrates a significant net gain in probability over comparable methods. Since the MC procedure was constructed to maximize the difference between the jth treatment mean and the control mean over all linear combinations, with the assumption of monotonicity, it performs well for all types of response curves and is recommended in identifying the MED.

Table 5.1

Probabilities of Identifying True MED with $k = 6$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, 1, 2, 3, 4, 5)$	2.5	4	.0002	.0126	.0313	.0481	.0538	.0342
		FWE	.0000	.0053	.0084	.0089	.0097	.0064
	2.0	4	.0010	.0346	.0770	.0977	.1051	.0727
		FWE	.0000	.0195	.0257	.0260	.0287	.0218
	1.5	3	.0005	.0306	.0540	.0486	.0523	.0397
		FWE	.0000	.0167	.0064	.0087	.0113	.0114
	1.0	3	.0044	.0650	.1141	.0968	.1024	.0832
		FWE	.0003	.0389	.0205	.0221	.0276	.0290
$(0, 0, 0, 0, 0, 5)$	2.5	6	.7406	.0010	.0850	.5541	.5107	.5649
	2.0	6	.8637	.0042	.2167	.6842	.6425	.7223
	1.5	6	.9379	.0162	.3925	.7942	.7582	.8524
	1.0	6	.9704	.0454	.6064	.8749	.8541	.9249
	0.0	FWE	.0490	.0464	.0517	.0495	.0491	.0474
$(0, 0, 0, 0, 5, 5)$	2.5	5	.2811	.0024	.1954	.3891	.4253	.5432
	2.0	5	.4854	.0102	.3474	.5370	.5789	.7178
	1.5	5	.6895	.0313	.5430	.6910	.7235	.8530
	1.0	5	.8368	.0734	.7042	.7928	.8195	.9196
	0.0	FWE	.0459	.0506	.0484	.0502	.0514	.0501
$(0, 0, 0, 5, 5, 5)$	2.5	4	.0245	.0085	.2904	.3014	.3961	.5326
	2.0	4	.0915	.0240	.4584	.4502	.5535	.7127
	1.5	4	.2308	.0639	.6348	.5982	.6964	.8396
	1.0	4	.4293	.1373	.7913	.7443	.8170	.9177
	0.0	FWE	.0443	.0485	.0498	.0496	.0496	.0498
$(0, 0.5, 5.5, 5)$	2.5	3	.0002	.0482	.3698	.2351	.3639	.4976
	2.0	3	.0014	.1081	.5674	.3904	.5386	.6745
	1.5	3	.0114	.2008	.7278	.5436	.6882	.8132
	1.0	3	.0576	.3308	.8471	.6942	.8110	.9022
	0.0	FWE	.0192	.0511	.0511	.0511	.0511	.0511
$(0, 5, 5, 5.5, 5)$	2.5	2	.0000	.4912	.0341	.2057	.4912	.3700
	2.0	2	.0000	.6398	.1077	.3406	.5416	.5358
	1.5	2	.0000	.7752	.2561	.5149	.7063	.7028
	1.0	2	.0001	.8662	.4445	.6572	.8234	.8203

Table 5.2

Probabilities of Identifying True MED with $k = 5$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, 1.25, 2.5, 3.75, 5)$	2.5	4	.0069	.0344	.0976	.1506	.1438	.0971
		FWE	.0003	.0161	.0286	.0302	.0306	.0209
	2.0	3	.0011	.0267	.0647	.0608	.0612	.0429
		FWE	.0001	.0106	.0044	.0055	.0066	.0069
	1.5	3	.0090	.0597	.1237	.1157	.1160	.0903
		FWE	.0003	.0293	.0176	.0192	.0226	.0230
	1.0	2	.0027	.0640	.0448	.0463	.0519	.0528
$(0, 0, 0, 0, 5)$	2.5	5	.7176	.0028	.1981	.5460	.5012	.5567
	2.0	5	.8526	.0199	.3535	.6773	.6360	.7320
	1.5	5	.9294	.0282	.5331	.7933	.7629	.8494
	1.0	5	.9652	.0654	.7162	.8773	.8548	.9241
	0.0	FWE	.0529	.0511	.0501	.0474	.0502	.0512
$(0, 0, 0, 5, 5)$	2.5	4	.2043	.0098	.2885	.3759	.4241	.5381
	2.0	4	.3906	.0255	.4675	.5297	.5783	.7215
	1.5	4	.5931	.0621	.6319	.6800	.7210	.8367
	1.0	4	.7617	.2358	.7800	.7963	.8270	.9232
	0.0	FWE	.0477	.0486	.0500	.0500	.0501	.0501
$(0, 0, 5, 5, 5)$	2.5	3	.0050	.0448	.4461	.2998	.3970	.5140
	2.0	3	.0267	.1011	.6137	.4459	.5570	.6851
	1.5	3	.1013	.1984	.7545	.6018	.7006	.8161
	1.0	3	.2258	.3318	.8609	.7426	.8215	.9050
	0.0	FWE	.0319	.0502	.0501	.0501	.0501	.0501
$(0, 5, 5, 5, 5)$	2.5	2	.0000	.4936	.0962	.2416	.3899	.3884
	2.0	2	.0000	.6500	.2159	.3980	.5648	.5631
	1.5	2	.0003	.7637	.3704	.5421	.7047	.7030
	1.0	2	.0030	.8659	.5758	.7023	.8317	.8311

Table 5.3

Probabilities of Identifying True MED with $k = 4$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, \frac{5}{3}, \frac{10}{3}, 5)$	2.5	3	.0103	.0376	.1135	.1127	.1070	.0745
		FWE	.0002	.0114	.0071	.0075	.0080	.0081
	2.0	3	.0319	.0079	.1975	.1927	.1831	.1403
		FWE	.0009	.0252	.0177	.0187	.0205	.0209
	1.5	2	.0056	.0610	.0469	.0482	.0519	.0526
	1.0	2	.0286	.1209	.1015	.1017	.1092	.1104
$(0, 0, 0, 5)$	2.5	4	.6952	.0075	.2974	.5491	.5103	.5764
	2.0	4	.8315	.0262	.4693	.6817	.6493	.7304
	1.5	4	.9142	.0716	.6501	.7944	.7676	.8525
	1.0	4	.9610	.1417	.7905	.8763	.8579	.9223
	0.0	FWE	.0483	.0501	.0465	.0465	.0474	.0474
$(0, 0, 5, 5)$	2.5	3	.1151	.0512	.4759	.3727	.4203	.5251
	2.0	3	.2358	.1119	.6419	.5322	.5904	.6914
	1.5	3	.4259	.1995	.7687	.6758	.7225	.8266
	1.0	3	.6157	.3289	.8724	.8044	.8393	.9099
	0.0	FWE	.0412	.0465	.0465	.0465	.0465	.0465
$(0, 5, 5, 5)$	2.5	2	.0001	.4972	.1834	.2903	.4102	.4135
	2.0	2	.0022	.6418	.3325	.4425	.5678	.5703
	1.5	2	.0105	.7784	.5164	.6080	.7241	.7263
	1.0	2	.0449	.8682	.6837	.7474	.8380	.8393

Table 5.4

Probabilities of Identifying at least one Effective Dose with $k = 6$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, 1, 2, 3, 4, 5)$	2.5	4	.1180	.1130	.4330	.5438	.5250	.4202
	2.0	4	.2359	.2346	.6559	.6874	.6817	.6144
	1.5	3	.3899	.3807	.8149	.7901	.7879	.7635
	1.0	3	.5624	.5673	.9257	.8799	.8863	.8894
$(0, 0, 0, 0. 0, 5)$	2.5	6	.7406	.0010	.0850	.5541	.5107	.5649
	2.0	6	.8637	.0042	.2167	.6842	.6425	.7223
	1.5	6	.9379	.0162	.3925	.7942	.7582	.8524
	1.0	6	.9704	.0454	.6064	.8749	.8541	.9249
$(0, 0, 0, 0, 5, 5)$	2.5	5	.3885	.0163	.6692	.5536	.5881	.7915
	2.0	5	.5720	.0472	.8341	.6856	.7344	.9126
	1.5	5	.7377	.1166	.9392	.7995	.8435	.9714
	1.0	5	.8614	.2309	.9741	.8748	.9092	.9881
$(0, 0, 0, 5, 5, 5)$	2.5	4	.1182	.1228	.8483	.5567	.6277	.8409
	2.0	4	.2358	.2362	.9439	.6852	.7651	.9455
	1.5	4	.3916	.3936	.9789	.7918	.8686	.9819
	1.0	4	.5614	.5668	.9845	.8754	.9318	.9888
$(0, 0, 5, 5, 5, 5)$	2.5	3	.0191	.3856	.6714	.5422	.6308	.7917
	2.0	3	.0502	.5721	.8369	.6774	.7778	.9157
	1.5	3	.1210	.7313	.9333	.7943	.8820	.9714
	1.0	3	.2335	.8565	.9725	.8728	.9410	.9843
$(0, 5, 5, 5, 5, 5)$	2.5	2	.0011	.7350	.0853	.5499	.6554	.5594
	2.0	2	.0048	.8646	.2104	.6843	.7970	.7272
	1.5	2	.0182	.9394	.4002	.7993	.8923	.8513
	1.0	2	.0495	.9775	.6024	.8776	.9509	.9280

Table 5.5

Probabilities of Identifying at least one Effective Dose with $k = 5$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, 1.25, 2.5, 3.75, 5)$	2.5	4	.1448	.1318	.4713	.5458	.5216	.4128
	2.0	3	.2643	.2503	.6556	.6802	.6613	.5803
	1.5	3	.4284	.3967	.7961	.7759	.7689	.7276
	1.0	2	.5996	.6003	.9083	.8817	.8848	.8679
$(0, 0, 0, 0, 5)$	2.5	5	.7176	.0028	.1981	.5460	.5012	.5567
	2.0	5	.8526	.0199	.3535	.6773	.6360	.7320
	1.5	5	.9294	.0282	.5331	.7933	.7629	.8494
	1.0	5	.9652	.0654	.7162	.8773	.8548	.9241
$(0, 0, 0.5, 5)$	2.5	4	.2935	.0525	.7614	.5424	.5874	.7625
	2.0	4	.4673	.1158	.8867	.6746	.7271	.8968
	1.5	4	.6497	.2273	.9492	.7960	.8434	.9614
	1.0	4	.7908	.3708	.9782	.8743	.9127	.9869
$(0, 0.5, 5, 5)$	2.5	3	.0484	.2992	.7680	.5482	.6249	.7741
	2.0	3	.1171	.4657	.8891	.6847	.7718	.9001
	1.5	3	.2290	.6365	.9497	.7885	.8664	.9614
	1.0	3	.3761	.7875	.9770	.8746	.9329	.9815
$(0, 5, 5, 5, 5)$	2.5	2	.0033	.7171	.1941	.5423	.6396	.5636
	2.0	2	.0105	.8541	.3543	.6894	.7909	.7336
	1.5	2	.0301	.9312	.5348	.7963	.8855	.8492
	1.0	2	.0760	.9708	.7171	.8798	.9466	.9278

Table 5.6

Probabilities of Identifying at least one Effective Dose with $k = 4$ dose levels

			Method					
Case	δ	MED	HC	RH	LC	DR	W(1)	MC
$(0, \frac{5}{3}, \frac{10}{3}, 5)$	2.5	3	.1835	.1682	.4329	.5460	.4452	.4069
	2.0	3	.3122	.2857	.5915	.6648	.6403	.5516
	1.5	2	.4799	.4758	.7725	.7988	.7820	.7241
	1.0	2	.6466	.6422	.8798	.8802	.8729	.8428
$(0, 0, 0, 5)$	2.5	4	.6952	.0075	.2974	.5491	.5103	.5764
	2.0	4	.8315	.0262	.4693	.6817	.6493	.7304
	1.5	4	.9142	.0716	.6501	.7944	.7676	.8525
	1.0	4	.9610	.1417	.9223	.8763	.8579	.9223
$(0, 0, 5, 5)$	2.5	3	.1737	.1735	.7158	.5427	.5872	.7131
	2.0	3	.3033	.3157	.8465	.6751	.7351	.8490
	1.5	3	.4839	.4755	.9311	.7943	.8471	.9405
	1.0	3	.6475	.6427	.9692	.8802	.9213	.9729
$(0, 5, 5, 5)$	2.5	2	.0087	.6974	.2973	.5424	.6293	.5766
	2.0	2	.0256	.8298	.4666	.6777	.7659	.7306
	1.5	2	.0652	.9180	.6391	.7956	.8757	.8510
	1.0	2	.1426	.9687	.7874	.8808	.9411	.9289

Chapter 6

Summary and Suggestions for Future Work

In dose-response studies, methods of hypothesis testing are often used to determine which dose levels, if any, are effective and in particular, to identify the minimum effective dose (MED). A step-down testing procedure is often performed as inferences are given in a specific order to verify if a treatment differs significantly from a control level. Failure to achieve the desired inference at any step eliminates the need for further comparisons. Simultaneous confidence intervals are preferred over point estimates and test statistics as they quantify the difference between any treatment and control mean at any step of the testing

procedure. The optimization Theorem and Algorithm developed in this thesis have enabled efficient calculations of the optimal confidence lower bound for the mean difference under the assumption of monotonicity. The innovative Multiple Contrast procedure has demonstrated asymptotic equivalence to the likelihood ratio test under ordered restrictions and, through simulation, was found to be more powerful that competitive procedures for various dose-response shapes.

The optimization theorem outlined in this thesis may be applied to other inference problems with order restrictions. In particular, other aspects of dose-response and toxicity studies may be improved through such an optimization procedure. As an example, recent articles by Tamhane, Dunnett, Green and Wetherington (2001) and Hothorn and Hauschke (2000) have discussed identification of the maximum safe dose (MAXSD), which is important in both randomized clinical dose-finding studies for the safety endpoint and toxicological studies. These authors proposed multiple testing procedures for equivalence with *a priori* ordered contrasts where an acceptable risk, denoted δ is defined in advance.

Furthermore, as the assumptions of normality and homoskedasticity are rarely met in practice, nonparametric procedures should also be compared in any future simulation study. As noted by various authors (e.g. Ruberg (1995a) and Chuang-Stein and Agresti(1997)), dose-response methods for categorical vari-

ables are also implemented in clinical trials. Thus, it is of interest to determine if generalizations of the optimization method would result in more accurate and efficient inference procedures in the aforementioned cases.

References

Ayer, M., Brunk, H.D., Ewing, W.T., Reid, E.S. (1955) An Empirical Distribution Function for Sampling with Incomplete Information. *The Annals of Mathematical Statistics*, 26 (4), 641-647.

Bartholomew, D.J. (1959a) A Test of Homogeneity for Ordered Alternatives. *Biometrika*, 46 (1/2) 36-48.

Bartholomew, D.J. (1959b) A Test of Homogeneity for Ordered Alternatives II. *Biometrika*, 46 (3/4) 328-335.

Bartholomew, D.J. (1961a) A Test of Homogeneity of Means Under Restricted Alternatives. *Journal of the Royal Statistical Society, Series B*, 23 (2) 239-281.

Bartholomew, D.J. (1961b) Ordered Tests in the Analysis of Variance. *Biometrika*, 48(3/4) 325-332.

Bauer, P. (1997) A Note on Multiple Testing Procedures in Dose Finding. *Biometrics* 53, 1125-1128.

Bretz, F. and Hothorn, L. A. (2000) A powerful alternative to Williams' test with application to toxicological dose-response relationships of normally distributed data. *Environmental and Ecological Statistics* 7, 135-154.

Chen, Y.I. (1999) Nonparametric Identification of the Minimum Effective Dose. *Biometrics* 55, 1236-1240.

Chen, Y.I. and Jan, S.L. (2002) Nonparametric Identification of the Minimum Effective Dose for Randomized Block Designs. *Communications in Statistics - Simulation and Computation* 31(2), 301-312.

Chen, Y.I. and Wolfe, D.A. (1993) Nonparametric Procedures for Comparing Umbrella Pattern Treatment Effects with a Control in a One-Way layout. *Biometrics*, 49, 455-465.

Chuang-Stein, C. and Agresti, A. (1997) Tutorial in Biostatistics: A Review of Tests for Detecting a Monotone Dose-Response Relationship with Ordinal Response Data. *Statistics in Medicine* 16, 2599-2618.

Daniel, Wayne W. (1990) *Applied Nonparametric Statistics*, 2nd edition, Pacific Grove: Duxbury.

Davidian M, Carroll RJ, Smith W. (1988) Variance functions and the minimum detectable concentration in assays. *Biometrika* 75: 549-556.

Davis, K.A. (2002) Step-Down Simultaneous Confidence Interval Procedure for Dose-Response Studies. *B.Sc. Honours Dissertation*, Memorial University of Newfoundland.

Dawson, D. and Magee, L. (2001) The National Hockey League Entry Draft, 1969-1995: An Application of a Weighted Pool-Adjacent Violators Algorithm. *The American Statistician*, 55(3) 194-199.

Dunnett, C.W. (1955) A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *Journal of the American Statistical Association*, 50, 1096-1121.

Dunnett, C. W. and Tamhane, A. C. (1998) Some New Multiple–Test Procedures for Dose Finding. *Journal of Biopharmaceutical Statistics*, 8(3) 353–366.

Everitt, B. S. (2002) *Statistical Analyses using S-Plus*, 2nd edition. Boca Raton: Chapman & Hall/CRC.

Fligner, M.A. and Wolfe, D.A. (1982) Distribution-free tests for comparing several treatments with a control. *Statistica Neederlandica*, 36, 119-127.

Hamasaki, T., Isomura, T., Baba, M. and Goto, M. (2000) Statistical Approaches to Detecting Dose-Response Relationships. *Drug Information Journal* 34, 579-590.

Hochberg, Y. and Tamhane, A. C. (1987) *Multiple Comparison Procedures*, New York: John Wiley.

Hogg, R.V. (1965) On Models and Hypotheses with Restricted Alternatives. *Journal of the American Statistical Association* 60, 1153-1162.

Hothorn, L. A. and Bretz, F. (2000). One-sided Simultaneous Confidence Intervals for Effective Dose Steps in Unbalanced Designs. *Biometrical Journal* 42(8), 995-1006.

Hothorn, L. A. and Hauschke, D. (2000). Identifying the Maximum Safe Dose: A Multiple Testing Procedure. *Journal of Biopharmaceutical Statistics*, 10(1),

pp. 15-30.

Hsu, J. C. and Berger, R. L. (1999) Stepwise Confidence Intervals Without Multiplicity Adjustment for Dose–Response and Toxicity Studies. *Journal of the American Statistical Association,* 94, 468–482.

Jonckheere, A.R. (1954) A distribution-free k-sample test against ordered alternatives. *Biometrika* 41, 133-145.

Källén, A. and Larsson, P. (1999) Dose Response Studies: How do we make them conclusive? *Statistics in Medicine* 18, 629-641.

Kruskal, W.H., Wallis, W.A. (1952) Use of Ranks in One-Criterion Variance Analysis. *Journal of the American Statistical Association*, 47, 583-621.

Kuhn, H.W. and Tucker, A.W. (1951) Nonlinear Programming. *2nd Berkeley Symposium, Math. Statist. Probab. (University of California Press, Berkeley)*, 481–492.

Lee, C.I.C. (1996) On Estimation for Monotone Dose-Response Curves. *Journal of the American Statistical Association* 91 No. 435 1110-1119.

Lee, C.I.C., Peng, J. and Liu, L. (2002) Statistical Inference for Best Treatment versus a Control. *Technical Report*, Memorial University of Newfoundland.

Liu, L. (2001) *Simultaneous Statistical Inference for Monotone Dose–Response Means. Doctoral Dissertation*, Memorial University of Newfoundland.

Mann, H.B., Whitney, D.R. (1947) On a test of whether one of the two random variables is stochastically larger than the other. *Annals of Mathematical Statistics*, 18, 50-60.

Marcus, R. (1976) The powers of some tests of the equality of normal means against an ordered alternative. *Biometrika* 63, 177-183.

Marcus, R. and Peritz, E. (1976) Some Simultaneous Confidence Bounds in Normal Models with Restricted Alternatives. *Journal of the Royal Statistical Society, Ser. B*, 38. 157–165.

Marcus, R., Peritz, E. and Gabriel, K.R. (1976) On closed testing procedures with special reference to ordered analysis of variance. *Biometrika* 63, 655-660.

Naik, U.D. (1975) Some Selection Rules for Comparing p Processes With a Standard. *Communications in Statistics, Part A - Theory and Methods*, 4, 519 - 535.

Neuhäuser, M. (2002) Nonparametric Identification of the Minimum Effective Dose. *Drug Information Journal* 36, 881-888.

Neuhäuser, M., Liu, P.Y. and Hothorn, L. (1998) Nonparametric Tests for Trend: Jonckheere's Test, a Modification and a Maximum Test. *Biometrical Journal* 40, 899-909.

Peng, J. (2002). Statistical Inference for Treatment versus a Control. *Doctoral Dissertation*. Memorial University of Newfoundland.

Robertson, T., Wright, F.T. and Dykstra, R.L. (1988). *Order Restricted Statistical Inference*, Chichester, U.K.: John Wiley.

Rom, D.M., Costello, R.J. and Connell, L.T. (1994) On closed test procedures for dose-response analysis. *Statistics in Medicine* 13, 1583-1596.

Ruberg, S.J. (1989) Contrasts for identifying the minimum effective dose. *Journal of the American Statistical Association* 84, 816-822.

Ruberg, S.J. (1995a) Dose Response Studies. I. Some Design Considerations. *Journal of Biopharmaceutical Statistics*, 5(1), 1-14.

Ruberg, S.J. (1995b) Dose Response Studies. II. Analysis and Interpretation. *Journal of Biopharmaceutical Statistics*, 5(1). 15-42.

Shirley, E. (1977) A nonparametric equivalent of Williams' test for contrasting increasing dose levels of a treatment. *Biometrics*, 33, 386-389.

Spector, P. (1994) *An Introduction to S and S-Plus*. Belmont: Duxbury.

Tamhane, A.C. and Dunnett, C.W. (1999) Stepwise multiple test procedures with biometric applications. *Journal of Statistical Planning and Inference* 82, 55-68.

Tamhane, A.J., Dunnett, C., Green, J.W. and Wetherington, J.D. (2001) Multiple Test Procedures for Identifying the Maximum Safe Dose. *Journal of the American Statistical Association*, Vol. 96, No. 455, pp. 835 - 843.

Tamhane, A.C., Hochberg, Y. and Dunnett, C. (1996). Multiple Test Procedures for Dose Finding. *Biometrics* 52, 21-37.

Tao, J., Shi, N.Z., Guo, J. and Gao, W. (2002) Stepwise procedures for the identification of minimum effective dose with unknown variances. *Statistics and Probability Letters* 57, 121-131.

Terpstra, T.J. (1952) The Asymptotic Normality and Consistency of Kendall's Test Against Trend when Ties are Present. *Indagationes Mathematics*, 14, 327-333.

Tsai, K.T. (2000) Robust Williams Trend Test. *Communications in Statistics - Theory and Methods* 29 (5&6), 1327-1346.

Venables, W.N. and Ripley, B.D. (1999) *Modern Applied Statistics with S-Plus*, 3rd edition. New York: Springer Verlag.

Williams, D.A. (1971) A test for treatment differences when several dose levels are compared with a zero-dose control. *Biometrics*, 27, 103-117.

Williams, D.A. (1972) The comparison of several dose levels with a zero-dose control. *Biometrics*, 28, 519-531.

Williams, D.A. (1977) Some inference procedures for monotonically ordered normal means. *Biometrika* 64, 9-14.

Williams, D.A. (1986) A note on Shirley's Nonparametric Test for Comparing Several Dose Levels with a Zero-dose Control. *Biometrics* 42, 183-186.

Wright, T.A. (1988) The One-Way Analysis of Variance with Ordered Alternatives: A Modification of Bartholomew's $\bar{E}^2$ Test. *The Canadian Journal of Statistics*. 1. 75-85.

Appendix I - S-Plus code for Iterative Algorithm

```
# S-Plus program to compute optimal simultaneous lower bound (SCLB)
# Author: Karelyn Davis
# Date: August 3, 2003

# Program to compute MLE for simple order:

# S-Plus program to compute MLE for simple order case
# Author: Karelyn Davis
# Date: July 28th, 2003

pava.prg <- function(k,y,w) {

ys <- rep(0,k); ws <- rep(0,k); ys1<-0; ws1<-0; u<-rep(0,k)
for(j in 1:k) {
if (j==1) {
  ys1 <- y[1]*w[1]
  ws1 <- w[1]
  ys[1]_ys1; ws[1]_ws1
}
else {
  a <- ys[j-1] + y[j]*w[j]
```

```
  ys[j]_a
  b <- ws[j-1] + w[j]
  ws[j]_b
}
}
i1 <- 0; i0 <-0
while(i1 <k) {
if (i0==0) {
  i01 <- i0 + 1; i1 <- i01
  ymin <- (ys[i01])/(ws[i01])
    for(i in i01:k) {
avg <- (ys[i])/(ws[i])
  if (avg <= ymin){
  ymin <- avg
  i1 <- i}
else {break}
    }
}
else {
  i01 <- i0+1; i1 <- i01
  ymin <- (ys[i01]-ys[i0])/(ws[i01]-ws[i0])
for(i in i01:k) {
  avg <- (ys[i]-ys[i0])/(ws[i]-ws[i0])
  if (avg <= ymin) {
    ymin <- avg
    i1 <- i}
  else {break}
    }
}
for(j in i01:i1) {
u[j]_ymin
  }
i0 <- i1
}
list(mle=u)
    }
```

```
# Program to compute SCLB using algorithm

sclb.prg <- function(k, w, y, stdev, crit) {

c <- rep(0,k)
# MLE stored as u
u <- pava.prg(k, y, w)$mle
if (k==2){
   c[1]_(-1/w[1]); c[2]_(1/w[2])
}
else {
# Compute muhat
su <- 0; sw <- 0
for (i in 1:k) {
   su <- u[i]*w[i] + su
   sw <- w[i] + sw
}
muhat <- su/sw
# Find initial p, q
for (i in 1:k) {
   if(u[i] < muhat) p<- i
}
for (i in k:1) {
   if (u[i] > muhat) q <- i
}
# Find optimal p, q using algorithm
d <- 1
while(d > 0) {
np <- 0; ybp <- 0
for (i in 1:p) {
np <- np + w[i]
ybp <- u[i]*w[i] + ybp
}
ybp <- ybp/np
s1p <- 0
for (i in 1:p) {
s1p <- s1p + w[i]*(u[i]-ybp)^2
```

```
}
nq <- 0; ybq <- 0
for (j in q:k) {
nq <- nq + w[j]
ybq <- u[j]*w[j] + ybq
}
ybq <- ybq/nq
sqk <- 0
for (i in q:k) {
sqk <- sqk + w[i]*(u[i]-ybq)^2
}
# Find beta

Betap <- np*(u[p]-ybp)
Betaq <- nq*(ybq-u[q])
Betapq <- c(Betap, Betaq)
print(Betapq)
Beta <- max(Betapq)
# Compute Talpha

t1 <- s1p + sqk + (1/np + 1/nq)*Beta^2
talpha <- sqrt(t1)/stdev
print(talpha)
d <- talpha - crit
if (talpha < crit) {break}
else{
if (Betap > Betaq) {
  q <- q; m1 <- u[p]
  for (j in 1:(p-1)) {
if (u[j] < m1) {
 p <- j}
  }
}
else{
 p <- p; m2 <- u[q]
 for (j in k:(q+1)) {
if (u[j] > m2) {
```

```
q <- j}
 }
}
}
}
# Compute B^2

bsq <- (stdev^2 * crit^2 - s1p - sqk)/(1/np + 1/nq)
b <- sqrt(bsq)
# Find optimal coefficients, c:
for (i in 1:p) {
  a <- (-1/np) + (u[i] - ybp)/b
  c[i]_a
}
for (i in q:k) {
  b1 <- 1/nq + (u[i]-ybq)/b
  c[i]_b1
 }
}
# Compute SCLB
sc2 <- 0; scmu <- 0
for (i in 1:k) {
  sc2 <- sc2 + w[i]*(c[i])^2
  scmu <- scmu + w[i]*c[i]*u[i]
}
# Compute Optimal Simultaneous Lower Bound
sclb <- scmu - (crit*stdev*sqrt(sc2))
list(k = k, optimalcoeff = c, optimalsclb = sclb)
    }
```